



**Joana dos Santos
Ferreira**

**Modelação estatística aplicada à gestão hospitalar:
o caso do reinternamento**

**Statistical modeling applied to hospital
management: the case of hospital readmission**



**Joana dos Santos
Ferreira**

**Modelação estatística aplicada à gestão hospitalar:
o caso do reinternamento**

**Statistical modeling applied to hospital
management: the case of hospital readmission**

Relatório de estágio apresentado à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Matemática e Aplicações, área de especialização Estatística e Otimização, realizada sob a orientação científica da Doutora Adelaide Freitas, Professora Auxiliar do Departamento de Matemática da Universidade de Aveiro.

Aos meus avós.

o júri / the jury

presidente / president

Prof. Doutor Pedro Filipe Pessoa Macedo

Professor Auxiliar da Universidade de Aveiro

vogais / examiners committee

Prof. Doutora Maria Cristina Souto de Miranda

Professora Adjunta da Universidade de Aveiro

Prof. Doutora Adelaide de Fátima Baptista Valente Freitas

Professora Auxiliar da Universidade de Aveiro

**agradecimentos /
acknowledgements**

Quero agradecer aos meus pais por tudo o que fizeram por mim até hoje, pelo apoio incondicional e por me deixarem traçar o meu próprio caminho.

Ao Pedro por todo o apoio e por ter sempre as palavras certas, mesmo que não sejam as que eu quero ouvir.

À Karem, à Mariana e ao André por serem os melhores irmãos que poderia pedir.

Aos meus amigos por todos os momentos e por estarem sempre presentes mesmo que ausentes.

À minha orientadora, professora Adelaide Freitas, por toda a preocupação, pela partilha do seu conhecimento, pelo apoio e por me ajudar a encontrar sempre a luzinha no fundo do túnel. Muito obrigada.

Quero ainda agradecer à equipa Prologica por me terem recebido da forma como o fizeram, pelo apoio, ajuda e crescimento que me proporcionaram.

Quero ainda agradecer ao João Oliveira por toda a ajuda disponibilizada assim como à Daniela Moleiro.

Palavras Chave

Financiamento Hospitalar, Reinternamento, Correlação, Regressão Logística.

Resumo

O financiamento hospitalar é feito de acordo com os critérios gerais definidos num contrato-programa. Este contempla o que é pago por cada serviço prestado e ainda as penalizações que são efetuadas a cada hospital que não cumpra determinados requisitos. Um dos critérios de penalização é a elevada taxa de reinternamentos. No âmbito do estágio realizado na Prologica, estabeleceu-se como objetivo o estudo do reinternamento, modelando-o de forma a ser possível prever qual o risco de um dado episódio ser reinternamento.

Com base em dados recolhidos num dado hospital do país, foi realizada uma análise de regressão logística múltipla. O modelo estimado permitiu identificar, dado um paciente ter sido admitido num hospital, que a ausência\presença de tumor, o destino após a sua alta, o tipo de admissão, os diagnósticos (principal e restantes) atribuídos e o primeiro procedimento clínico realizado nessa admissão são fatores que determinam significativamente o risco de tal admissão se tratar de um episódio de reinternamento.

Keywords

Hospital Financing, Hospital Readmission, Correlation, Logistic Regression.

Abstract

Hospital financing is done according to the general criteria defined in a contract program. Payments of services provided by the hospital but also penalties stemming from the fact that the hospital does not fulfill certain requirements are two aspects contemplated in the contract program. One of the penalty criteria is the high hospital readmission rate. In the scope of the internship at the company Prologica, hospital readmission was analyzed, in particular, a model to predict the risk of patient readmission was constructed. The results of this work allow us to conclude that the absence of tumor, the type of admission, the diagnostics performed and the first procedure performed are factors that significantly affect the risk of an episode being a readmission.

Conteúdo

Conteúdo	i
Lista de Figuras	iii
Lista de Tabelas	v
1 Introdução	1
1.1 Financiamento Hospitalar	1
1.2 Contrato-Programa	3
1.3 Reinternamento	3
1.4 Objetivos do Relatório	4
1.5 Estrutura do Relatório	4
2 Medidas de Associação	5
2.1 Medidas de associação para dados categóricos	6
2.2 Medidas de associação entre dados categóricos e numéricos	7
3 Regressão Logística	11
3.1 Regressão Logística Simples	12
3.1.1 Método de Máxima Verossimilhança	13
3.1.2 Teste de Wald	14
3.1.3 Razão de chances (<i>Odds Ratio</i>)	14
3.2 Regressão Logística Múltipla	15
3.2.1 Estimação dos Parâmetros	16
3.2.2 Teste de Wald	16
3.2.3 Razão de chances ajustada (<i>Adjusted Odds Ratio</i>)	16
3.2.4 Seleção de variáveis preditoras	17
3.2.5 Critérios de Informação	17
3.2.6 Teste de Hosmer & Lemeshow	18
3.2.7 Curva ROC	18
4 Base de Dados	21

5	Resultados	29
5.1	Correlação	29
5.2	Regressão Logística Simples	30
5.3	Regressão Logística Múltipla	48
5.4	Comparação dos Modelos	51
6	Aplicação <i>Shiny</i>	55
7	Conclusões	59
	Referências	67
	Apêndice A	
	Tabelas relativas aos modelos construídos	69
	Apêndice B	
	Código utilizado no <i>Software RStudio</i>	117

Lista de Figuras

1.1	Comparação da produtividade hospitalar segundo o indicador da taxa anual de ocupação de internamento. Fonte: http://benchmarking.acss.min-saude.pt/BH_ProdutividadeDashboard	2
2.1	Relações entre variáveis aleatórias: linear, exponencial, logarítmica e quadrática. Fonte: Hall[4]	5
5.1	Curvas ROC e respetivo AUC dos Modelos 1 a 12	52
6.1	Exemplo de um caso com risco elevado de se tratar de um episódio de reinternamento . .	55
6.2	Exemplo de um caso com risco moderado de se tratar de um episódio de reinternamento	56
6.3	Exemplo de um caso com risco baixo de se tratar de um episódio de reinternamento . . .	56
6.4	Exemplo de mau preenchimento na aplicação <i>Shiny</i>	56

Lista de Tabelas

2.1	Exemplo de uma tabela de contingência que estuda se uma patologia está ou não relacionada com o tabagismo dos indivíduos. Fonte: Exemplo adaptado de Hall [4].	6
2.2	Dados que permitirão estudar a relação entre as notas de um curso e o medo de falhar. . .	8
2.3	Estudo da relação entre as notas de um curso e o medo de falhar	8
3.1	Matriz de confusão onde VN significa verdadeiros positivos, FN corresponde a falsos negativos, FP remete a falsos positivos e VP representa verdadeiros positivos	19
4.1	Descrição das variáveis restantes após primeira análise	21
4.2	Codificação da variável DIA_ADM	22
4.3	Codificação da variável ADM_TIP	23
4.4	Codificação das variáveis DIA_2, DIA_3, DIA_4 e DIA_5	24
4.5	Codificação da variável DSP	24
4.6	Codificação da variável MOT_TRANF	26
4.7	Codificação da variável RINT	26
4.8	Codificação da variável SEXO	26
4.9	Codificação das variáveis PROC_1, PROC_2, PROC_3, PROC_4 e PROC_5	27
4.10	Variáveis finais	28
5.1	Medidas de Associação	30
5.2	Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável Sexo	30
5.3	Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável MORF_TUM	31
5.4	Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável DSP	31
5.5	Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável ADM_TIP	33
5.6	Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável MOT_TRANF	33

5.7	Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável DIA_ADM	34
5.8	Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável DIA_1	36
5.9	Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável DIA_2	37
5.10	Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável DIA_3	39
5.11	Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável DIA_4	40
5.12	Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável DIA_5	41
5.13	Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável PROC_1	42
5.14	Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável PROC_2	43
5.15	Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável PROC_3	44
5.16	Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável PROC_4	45
5.17	Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável PROC_5	46
5.18	Estimativas associadas aos modelos de regressão logística simples tomando cada uma das variáveis numéricas	47
5.19	Exemplo ilustrativo da transformação utilizada	49
5.20	Ilustração da diferença entre a utilização de variáveis <i>dummy</i> e a transformação proposta	49
5.21	Medidas de Avaliação dos Modelos 1 a 6	51
5.22	Medidas de Avaliação dos Modelos 7 a 12	51
5.23	Número de variáveis e variáveis de cada modelo	53
7.1	Significância estatística das variáveis presentes no modelo selecionado: IC, intervalo de confiança; OR, Razão de chances (<i>Odds Ratio</i>); AOR, Razão de chances ajustada (<i>Adjusted Odds Ratio</i>).	60
1	Coeficientes β , Teste de Wald (p-valor), OR e respectivo IC a 95% relativos ao Modelo 1 .	70
2	Coeficientes β , Teste de Wald (p-valor), OR e respectivo IC a 95% relativos ao Modelo 2 .	76
3	Coeficientes β , Teste de Wald (p-valor), OR e respectivo IC a 95% relativos ao Modelo 3 .	80
4	Coeficientes β , Teste de Wald (p-valor), OR e respectivo IC a 95% relativos ao Modelo 4 .	84

5	Coeficientes β , Teste de Wald (p-valor), OR e respetivo IC a 95% relativos ao Modelo 5 .	87
6	Coeficientes β , Teste de Wald (p-valor), OR e respetivo IC a 95% relativos ao Modelo 6 .	93
7	Coeficientes β , Teste de Wald (p-valor), OR e respetivo IC a 95% relativos ao Modelo 7 .	96
8	Coeficientes β , Teste de Wald (p-valor), OR e respetivo IC a 95% relativos ao Modelo 8 .	102
9	Coeficientes β , Teste de Wald (p-valor), OR e respetivo IC a 95% relativos ao Modelo 9 .	104
10	Coeficientes β , Teste de Wald (p-valor), OR e respetivo IC a 95% relativos ao Modelo 10 .	106
11	Coeficientes β , Teste de Wald (p-valor), OR e respetivo IC a 95% relativos ao Modelo 11 .	109
12	Coeficientes β , Teste de Wald (p-valor), OR e respetivo IC a 95% relativos ao Modelo 12 .	115

Introdução

1.1 FINANCIAMENTO HOSPITALAR

Em Portugal, o Serviço Nacional de Saúde (SNS) foi estabelecido em 1976 pela Constituição da República Portuguesa. É financiado através do Orçamento Geral de Estado com base no modelo retrospectivo. Este sistema foi inspirado no modelo Beveridge que defende o direito à proteção da saúde de uma forma universal, geral e gratuita. O modelo Beveridge prevê o reembolso aos hospitais com base na despesa real do ano anterior, não incentivando a eficiência nem o controlo de gastos, já que quanto mais gastam mais recebem no ano seguinte, existindo assim um consumo exagerado dos recursos; [1].

Em 1984, adaptou-se o sistema de Grupos de Diagnóstico Homogéneo (GDH). Mas apenas em 1997 se deu a aplicação total do mesmo. Assim, os pagamentos a todos os hospitais do país combinavam a componente histórica e a componente de produção ajustada pelos GDH. Nos anos seguintes, o modelo de pagamento foi sendo substituído, gradualmente, pelo modelo prospetivo que tem em consideração estes grupos; [1].

O modelo prospetivo prevê reembolso dos hospitais com base numa quantia fixa, pré-estabelecida, por doente, sem considerar o que realmente o doente gastou. A utilização deste modelo estimula a eficiência e a poupança de recursos já que o hospital se torna responsável por possíveis prejuízos. Contudo, os doentes podem sair prejudicados, pois o seu tratamento pode não ser benéfico, em termos económicos, para o hospital; [1], [2].

Já em 2002 foi estabelecido um novo regime jurídico, passando a existir uma rede de prestação de cuidados de saúde, constituída por hospitais do Setor Público Administrativo (SPA), hospitais Entidade Pública Empresarial (EPE), hospitais Sociedade Anónima (SA), e estabelecimentos privados, com ou sem fins lucrativos. Com a criação desta nova rede, o SNS pôde complementar os serviços em falta, diminuir as listas de espera cirúrgicas, e melhorar outros problemas identificados; [1]–[3].

Segundo Silva (2015) [1], este novo regime foi imprescindível na reforma da gestão hospitalar e na melhoria do desempenho global do SNS, tendo por base quatro grandes princípios:

1. criação de contratos-programa;
2. agrupamento de cuidados e serviços por linhas de atividade (internamento, urgências,

- consultas externas e hospital de dia), com o custo ajustado pelo case-mix; ¹
3. contratação de um volume de produção que tenha em conta a capacidade instalada, permitindo cobrir os custos fixos do respetivo hospital;
 4. produção marginal remunerada até um certo limite médio.

Tornou-se possível comparar o desempenho entre os hospitais através de vários indicadores, o chamado *benchmarking*. No site: <http://benchmarking.acss.min-saude.pt> é possível fazer essa análise. Por exemplo, tornou-se possível comparar a produtividade hospitalar segundo o indicador da taxa anual de ocupação de internamento:

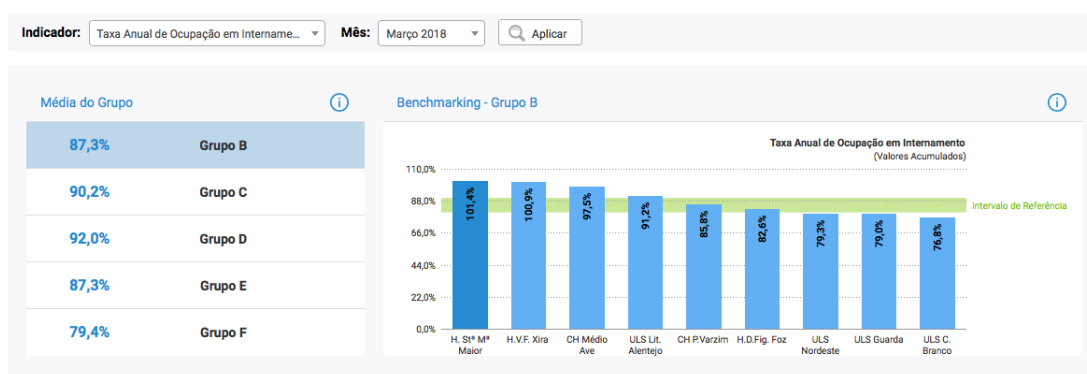


Figura 1.1: Comparação da produtividade hospitalar segundo o indicador da taxa anual de ocupação de internamento. Fonte: http://benchmarking.acss.min-saude.pt/BH_ProdutividadeDashboard

Ainda em 2002, foi estabelecido o programa Parcerias Público Privadas (PPP); [1], [3]. Segundo a Administração Central do Sistema de Saúde ², este programa foi criado "com vista a promover formas inovadoras de partilha do risco para a prestação de cuidados de saúde, como novas experiências de gestão, bem como a participação do setor privado na conceção, construção, financiamento e exploração de unidades hospitalares do Serviço Nacional de Saúde (SNS)."

¹ Coeficiente global de ponderação da produção que reflete a relatividade de um hospital face aos outros, em termos da sua maior ou menor proporção de doentes com patologias complexas e, consequentemente, mais consumidoras de recursos. Determina-se calculando o rácio entre o número de doentes equivalentes ponderados pelos pesos relativos dos respetivos GDH e o número total de doentes equivalentes. Fonte: http://www2.acss.min-saude.pt/Portals/0/Cálculo%20do%20doente%20equivalente%20e%20ICM_2009_Finaln.pdf

Os doentes equivalentes, num dado GDH, correspondem aos episódios de internamento que se obtêm após a transformação dos dias de internamento dos episódios de curta duração e dos doentes transferidos de cada GDH, em episódios equivalentes aos episódios típicos ou normais do respetivo GDH. Fonte: http://www2.acss.min-saude.pt/Portals/0/Cálculo%20do%20doente%20equivalente%20e%20ICM_2009_Finaln.pdf

² Fonte: <http://www.acss.min-saude.pt/2016/10/12/parcerias-publico-privadas/>

1.2 CONTRATO-PROGRAMA

O contrato-programa é realizado após a negociação entre a Administração Central do Sistema de Saúde (ACSS), os Departamentos de Contratualização de cada Administração Regional de Saúde (ARS) e cada unidade hospitalar. Têm uma vigência de 3 anos e devem ser revistos anualmente.

Este contrato estabelece as quantidades a contratar por linha de produção, assim como os limites à produção marginal que corresponde ao extra produzido pelo hospital, define a tabela de preços a praticar e os critérios de atribuição de incentivos. São também definidos os objetivos a atingir e ainda os indicadores que possibilitam avaliar o desempenho do hospital. O contrato-programa antecipa ainda reduções e/ou penalizações no pagamento caso não seja cumprido o contratado; [1], [2].

Citando Silva (2015), " A quantidade a contratar por linha de produção é calculada mediante o plano de desempenho, elaborado no final do exercício financeiro do ano anterior a que respeita o contrato-programa. Por cada linha de produção foi estabelecida uma fórmula de cálculo de pagamento. (...) Os contratos-programa estabelecem ainda outras componentes de financiamento, nomeadamente produção de linhas específicas consideradas prioritárias; definição do valor da percentagem a ser paga como incentivo, caso sejam cumpridas as metas negociadas; e determinação do pagamento a ser realizado, por linha de produção, caso seja produzido mais ou menos em relação ao volume contratado." [1, pág.13]

Os contratos-programa são disponibilizados no site da ACSS ³, e encontram-se divididos por regiões sendo possível analisar qualquer contrato-programa realizado.

Atualmente existem no SNS:

- 4 instituições em regime PPP.
- 7 instituições em regime SPA.
- 39 instituições em regime EPE.

1.3 REINTERNAMENTO

É considerado reinternamento caso um doente seja internado uma segunda vez na mesma unidade de saúde e o seu episódio esteja clinicamente relacionado com o anterior.

De acordo com a Circular Normativa nº 33/2012 de 19 de julho de 2012, não é considerado reinternamento os casos que:

- o segundo internamento não está clinicamente relacionado com o anterior, desde que assim demonstrado pela entidade prestadora, e as situações do foro oncológico.
- o segundo internamento acontece após saída contra parecer médico.
- o doente foi transferido para realização de exame que obrigue a internamento, seguindo-se o tratamento no hospital de origem.

Para este estudo serão considerados como reinternamento os casos que cumpram o segundo ponto, pois pode servir de decisão de apoio ao doente. Todos os reinternamentos considerados foram registados a 5 dias.

³<http://www.acss.min-saude.pt//category/cuidados-de-saude/hospitalares/>

1.4 OBJETIVOS DO RELATÓRIO

O presente relatório baseia-se no trabalho desenvolvido, enquanto estagiária, na empresa Prologica. Este trabalho teve como intuito modelar uma característica associada à gestão hospitalar, nomeadamente a taxa de reinternamento. Importa referir que elevadas taxas de reinternamento acarretam penalizações no financiamento hospitalar. O objetivo final do trabalho passou pelo desenvolvimento de uma aplicação *web* de fácil utilização que permitisse ao médico responsável calcular o risco de um determinado episódio ser reinternamento. Para tal, recorreu-se a modelos de regressão logística, de forma a identificar os fatores mais significativos na caracterização de um episódio de reinternamento. A análise de resíduos não faz parte dos objetivos do presente estudo.

1.5 ESTRUTURA DO RELATÓRIO

Este relatório contém sete capítulos. No primeiro capítulo é feita uma breve reflexão sobre os conceitos de financiamento hospitalar, contrato-programa e reinternamento, sendo também referidos os objetivos e a estrutura do relatório.

Nos capítulos 2 e 3 é feita uma exposição sobre medidas de correlação e sobre o modelo de regressão logística, respetivamente, por corresponder a conhecimentos que serão usados na aplicação apresentada neste relatório.

No capítulo 4 é exposta a base de dados com a qual se trabalhou. No capítulo 5 é exposto todo o processo de análise realizado, isto é, a análise de associação das variáveis, a seleção de variáveis preditoras, a obtenção dos modelos de regressão logística e ainda o processo de seleção do modelo final que conduziu à construção de uma aplicação *web* que é apresentada no capítulo 6. Por fim, encontra-se o capítulo 7 com as conclusões obtidas no decorrer deste trabalho assim como propostas para trabalhos futuros.

Medidas de Associação

Quando se faz um estudo estatístico é frequente perguntarmo-nos qual a relação existente entre as variáveis em estudo, caso essa exista. Contudo, nem sempre é fácil concluir se duas variáveis são ou não independentes entre si ou qual o grau de dependência existente entre quaisquer duas. A exposição que se segue foi escrita com base nas referências Hall[4] e Cramer (2003)[5].

Uma das formas mais usuais de associação entre variáveis numéricas é a linear. Entre duas variáveis, esta relação pode ser observável através de um gráfico de dispersão; se os pontos estiverem dispostos sobre uma reta, diz-se que existe linearidade perfeita. Outras relações não lineares poderão ser visíveis graficamente, como por exemplo, a exponencial, a logarítmica e a quadrática. (Figura 2.1)

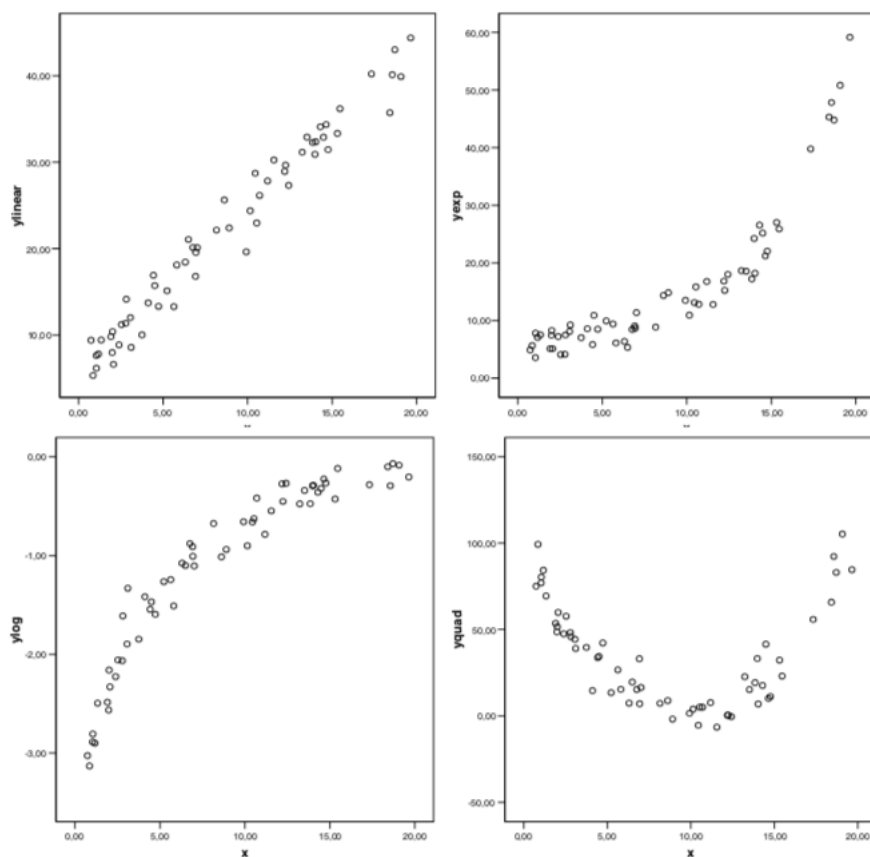


Figura 2.1: Relações entre variáveis aleatórias: linear, exponencial, logarítmica e quadrática. Fonte: Hall[4]

Para medir ou avaliar a associação entre variáveis existem alguns passos que devem ser considerados:

1. Construir diagramas de dispersão entre pares de variáveis. Nesse caso estamos perante uma das opções:
 - se duas variáveis são independentes, os pontos estarão dispostos aleatoriamente.
 - se for observável um conjunto de pontos que se dispõem aproximadamente sobre uma reta estamos perante uma relação linear, a qual poderá contudo não ser de fácil identificação.
2. Calcular medidas de associação: seleciona-se a medida de associação adequada ao tipo de variáveis a estudar.
 - Se as variáveis forem numéricas as medidas de associação designam-se por coeficientes de correlação e o coeficiente mais utilizado é o de Pearson.
 - Se as variáveis forem categóricas existem algumas opções baseadas na estatística de Pearson. A abordada aqui será a medida de associação de Cramér pois as variáveis encontram-se distribuídas por categorias nominais não ordenáveis.
 - Se uma variável for numérica e outra categórica, poderá ser utilizada a medida de associação eta, como se fará neste trabalho mais à frente.
3. Realizar um teste de hipóteses para determinar se os valores calculados com base na amostra no passo anterior levam a concluir que a associação na população é significativa ou não.

2.1 MEDIDAS DE ASSOCIAÇÃO PARA DADOS CATEGÓRICOS

Quando se trata de dados categóricos, as medidas para avaliar o grau de associação entre duas variáveis dependem do número de categorias das variáveis. Se existirem apenas duas categorias em cada variável, é possível visualizar os dados numa tabela com duas linhas e duas colunas que contêm as frequências observadas para cada combinação de categorias. A esta tabela dá-se o nome de tabela de contingência.

Imagine-se que se pretende estudar se uma patologia está ou não relacionada com o tabagismo dos indivíduos. Um exemplo de uma tabela de contingência, para uma amostra hipotética de 200 indivíduos, seria:

Tabagismo	Patologia		Total
	Presente	Ausente	
Fumador	90	40	130
Não Fumador	35	35	70
Total	125	75	200

Tabela 2.1: Exemplo de uma tabela de contingência que estuda se uma patologia está ou não relacionada com o tabagismo dos indivíduos. Fonte: Exemplo adaptado de Hall [4].

Teste do χ^2

Através da tabela de contingência é possível calcular uma estatística que permite fazer um teste de hipóteses denominado qui-quadrado, χ^2 , para avaliar a independência entre duas variáveis:

$$H_0 : \text{as variáveis são independentes} \quad vs \quad H_1 : \text{as variáveis são dependentes.}$$

A estatística do teste de qui-quadrado é definida por:

$$X^2 = \sum_{\text{todas as células}} \frac{(O_i - E_i)^2}{E_i}, \quad (2.1)$$

onde E_i representa a frequência esperada sob hipótese H_0 , que se obtém multiplicando as margens da tabela e dividindo pelo número total de observações e O_i representa a frequência observada.

O nome do teste é χ^2 pois, quando o número de observações é elevado, a distribuição da estatística X^2 é aproximadamente a do *qui-quadrado*.

A hipótese nula é rejeitada quando o valor observado da estatística de teste é superior ao quantil de ordem igual a um certo α pré-estabelecido da distribuição de *qui-quadrado* com $(r - 1)(c - 1)$ graus de liberdade, onde r corresponde ao número de linhas e c ao número de colunas da tabela de contingência.

Coefficiente de Cramér

O coeficiente de Cramér é obtido através da estatística X^2 apresentada anteriormente, e é dado por:

$$C = \sqrt{\frac{X^2}{n(l - 1)}} \quad (2.2)$$

onde n é o número total de observações e l é o mínimo entre o número de linhas e colunas da tabela de contingência, ($l = \min(r, c)$).

O valor deste coeficiente varia entre 0 e 1, sendo que:

- $C = 0$ corresponde a variáveis totalmente independentes.
- $C = 1$ poderá significar associação perfeita entre as variáveis. A associação é perfeita quando as duas variáveis têm o mesmo número de categorias, caso contrário são fortemente associadas.

2.2 MEDIDAS DE ASSOCIAÇÃO ENTRE DADOS CATEGÓRICOS E NUMÉRICOS

Coefficiente eta

O coeficiente eta é obtido através da seguinte expressão:

$$\eta = \sqrt{\frac{\text{soma de quadrados entre grupos}}{\text{soma de quadrados total}}}. \quad (2.3)$$

O valor deste coeficiente varia entre 0 e 1, sendo que:

- $\eta = 0$ corresponde a variáveis totalmente independentes.
- $\eta = 1$ significa associação perfeita entre as variáveis.

Uma forma de testar a significância desta medida de associação, isto é, de testar

$$H_0 : \eta = 0 \quad vs \quad H_1 : \eta \neq 0,$$

é usar a estatística do teste F dada pela razão:

$$F = \frac{\eta^2 / (k - 1)}{(1 - \eta^2) / (n - k)} \quad (2.4)$$

onde k é o número de categorias da variável nominal e n é o número total de observações.

De seguida apresenta-se um exemplo para ilustrar o cálculo do coeficiente eta tomando a expressão 2.3. O exemplo retirado de Cramer (2003) [5] é relativo ao estudo da existência de associação entre as notas de um curso e o medo de falhar:

Notas do curso	Medo de falhar
1	1
2	1
2	2
3	3
4	3
2	4
1	5

Tabela 2.2: Dados que permitirão estudar a relação entre as notas de um curso e o medo de falhar.

Agrupando os dados:

	Medo de falhar					
	1	2	3	4	5	
Notas do curso	1	2	3	2	1	
	2		4			
Soma:	3	2	7	2	1	Soma total:15
N	2	1	2	1	1	
Média	1.5	2	3.5	2	1	Média total: 2.1

Tabela 2.3: Estudo da relação entre as notas de um curso e o medo de falhar

Começa-se por calcular a soma dos quadrados entre grupos. Subtraí-se a média total à média do grupo, eleva-se ao quadrado e multiplica-se pelo número de casos registados em cada grupo,

somando todos os produtos obtidos. Para o exemplo em concreto, tem-se:

$$\begin{aligned}
 \text{soma de quadrados entre grupos} &= [(1.5 - 2.1)^2 \times 2] + [(2.0 - 2.1)^2 \times 1] + [(3.5 - 2.1)^2 \times 2] + \\
 &\quad [(2.0 - 2.1)^2 \times 1] + [(1.0 - 2.1)^2 \times 1] \\
 &= 0.72 + 0.00 + 3.92 + 0.00 + 1.21 \\
 &= 5.85.
 \end{aligned} \tag{2.5}$$

De seguida, calcula-se a soma dos quadrados total. Subtrai-se cada valor registado à média total, eleva-se ao quadrado e somam-se todas as parcelas. Concretizando no exemplo:

$$\begin{aligned}
 \text{soma de quadrados total} &= (2.1 - 1)^2 + (2.1 - 2)^2 + (2.1 - 2)^2 + (2.1 - 3)^2 \\
 &\quad (2.1 - 4)^2 + (2.1 - 2)^2 + (2.1 - 1)^2 \\
 &= 1.2 + 0.0 + 0.0 + 0.8 + 3.6 + 0.0 + 1.2 \\
 &= 6.8.
 \end{aligned} \tag{2.6}$$

Por fim, divide-se a soma de quadrados entre grupos pela soma de quadrados total e calcula-se a raíz quadrada do valor resultante. Aplicando ao exemplo:

$$\begin{aligned}
 \eta &= \sqrt{\frac{\text{soma de quadrados entre grupos}}{\text{soma de quadrados total}}} \\
 &= \sqrt{\frac{5.85}{6.8}} \\
 &= 0.928.
 \end{aligned} \tag{2.7}$$

Regressão Logística

Os modelos lineares generalizados, MLG, foram introduzidos em 1972 por Nelder e Wedderburn [8], com o intuito de colmatar as dificuldades sentidas na época, em modelar dados que não seguissem uma distribuição gaussiana que impossibilitavam, assim, o uso do modelo de regressão linear clássico. Por isso mesmo, os MLG são uma extensão deste e definem-se pela seguinte estrutura:

- Componente aleatória: corresponde à variável resposta Y . Esta variável deve ser aleatória com distribuição de probabilidade pertencente à família exponencial de distribuições.

Diz-se que uma variável aleatória Y tem distribuição na família exponencial se a sua função densidade de probabilidade ou função massa de probabilidade se pode escrever na forma

$$f(y \mid \theta, \phi) = \exp \left\{ \frac{y\theta - b(\theta)}{a(\theta)} + c(y, \phi) \right\}, \quad (3.1)$$

onde θ é o parâmetro natural, ϕ é um parâmetro de dispersão, em geral conhecido, e $a(\cdot)$, $b(\cdot)$ e $c(\cdot, \cdot)$ são funções reais conhecidas.

- Componente sistemática: consiste numa combinação linear de variáveis preditoras.

Considerando que existem p variáveis preditoras e n observações, e descrevendo essas observações numa matriz \mathbf{X} dada por:

$$\mathbf{X} = \begin{bmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_n^T \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix}$$

e dado um vetor de parâmetros

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix},$$

define-se a combinação linear $\eta_i = x_i^T \boldsymbol{\beta}$ como sendo o preditor linear, para cada i , $i = 1, \dots, n$ da resposta y_i .

- Função de ligação: é uma função diferenciável e monótona, g , que relaciona as componentes aleatória e sistemática e depende da distribuição de Y ,

$$\eta_i = g(\mu_i), i = 1, 2, \dots, n.$$

Ao longo deste capítulo será abordado o caso específico quando Y é uma variável binária, isto é, apenas toma dois valores, $Y_i = 0$ e $Y_i = 1$. Usualmente usa-se a designação de "insucesso" e "sucesso", respetivamente. Pode ser dito que Y segue uma distribuição de Bernoulli onde p é a probabilidade de sucesso numa experiência de Bernoulli. Verificando-se que a função de probabilidade de Y pode ser escrita por:

$$\begin{aligned} f(y; p) &= p^y(1-p)^{1-y} \\ &= p^y(1-p)(1-p)^{-y} \\ &= (1-p) \frac{p^y}{(1-p)^y} \\ &= (1-p) \left(\frac{p}{1-p} \right)^y \\ &= \exp \left\{ \log(1-p) + \log \left(\frac{p}{1-p} \right)^y \right\} \\ &= \exp \left\{ \log(1-p) + y \log \left(\frac{p}{1-p} \right) \right\} \end{aligned} \tag{3.2}$$

de acordo com a expressão 3.1, tem-se: $\theta = \log \left(\frac{p}{1-p} \right)$, $a(\theta) = 1$, $b(\theta) = \log(1 + e^\theta)$ e $c(y, \phi) = 0$. Ora vejamos:

$$\theta = \log \left(\frac{p}{1-p} \right) \Leftrightarrow e^\theta = \frac{p}{1-p} \Leftrightarrow (1-p)e^\theta = p \Leftrightarrow p = \frac{e^\theta}{1+e^\theta} \tag{3.3}$$

e

$$\begin{aligned} b(\theta) &= -\log(1-p) \\ &= -\log\left(1 - \frac{1-e^\theta}{1+e^\theta}\right) \\ &= -\log\left(\frac{1}{1+e^\theta}\right) \\ &= \log(1+e^\theta). \end{aligned} \tag{3.4}$$

Desta forma, fica demonstrado que Y , nas condições referidas, tem distribuição na família exponencial.

3.1 REGRESSÃO LOGÍSTICA SIMPLES

Considere-se x_i a variável preditora e y_i o número de vezes que dado acontecimento ocorre, sendo $i = 1, \dots, n$, com n igual ao número de observações. Considere-se ainda que Y_i segue uma distribuição de Bernoulli, e probabilidade de sucesso $\pi(x_i)$ com $E[Y] = \pi(x_i)$. Logo,

$$P[Y_i = y_i] = \pi(x_i)^{y_i}(1 - \pi(x_i))^{1-y_i}. \tag{3.5}$$

A função de regressão logística simples é dada por:

$$\pi(x_i) = \frac{\exp(\beta_0 + \beta_1 x_i)}{1 + \exp(\beta_0 + \beta_1 x_i)}, \quad i = 1, \dots, n. \quad (3.6)$$

A transformação *logit* - função de ligação para o modelo de regressão logística, denotado por $g(x)$, é dada por:

$$g(x) = \ln \left(\frac{\pi(x)}{1 - \pi(x)} \right) = \beta_0 + \beta_1 x. \quad (3.7)$$

3.1.1 Método de Máxima Verosimilhança

Para estimar os parâmetros β_0 e β_1 utiliza-se o método de máxima verosimilhança. Este método calcula os valores de β_0, β_1 que maximizam a função de verosimilhança dada uma amostra, sendo estas as estimativas dos parâmetros desconhecidos com base naquela amostra. No caso em estudo, com apenas dois resultados possíveis (sucesso $Y = 1$ e insucesso $Y = 0$), e assumindo a independência das observações, a função de verosimilhança é dada por:

$$l(\beta_0, \beta_1) = \prod_{i=1}^n \pi(x_i)^{y_i} (1 - \pi(x_i))^{1-y_i} \quad (3.8)$$

onde $\pi(x_i)$ representa a probabilidade de sucesso.

Frequentemente é utilizada a expressão do logaritmo da verosimilhança por ser mais fácil de manipular algebricamente:

$$\begin{aligned} L(\beta_0, \beta_1) &= \ln(l(\beta_0, \beta_1)) \\ &= \ln \left(\prod_{i=1}^n \pi(x_i)^{y_i} (1 - \pi(x_i))^{1-y_i} \right) \\ &= \sum_{i=1}^n \left(y_i \ln \left(\frac{\pi(x_i)}{1 - \pi(x_i)} \right) + \ln(1 - \pi(x_i)) \right) \end{aligned} \quad (3.9)$$

Substituindo $\pi(x_i)$ pela expressão (3.6) tem-se:

$$\begin{aligned} L(\beta_0, \beta_1) &= \ln(l(\beta_0, \beta_1)) \\ &= \sum_{i=1}^n y_i (\beta_0 + \beta_1 x_i) - \ln(1 + \exp(\beta_0 + \beta_1 x_i)). \end{aligned} \quad (3.10)$$

Como em qualquer problema de maximização, para obter o valor que maximiza $L(\beta_0, \beta_1)$ deriva-se (3.10) em ordem a cada parâmetro e igualam-se as equações de verosimilhança a zero. Concretamente, resulta:

$$\begin{aligned} \frac{\partial L(\beta_0, \beta_1)}{\partial \beta_0} &= \sum_{i=1}^n \left(y_i - \frac{\exp(\beta_0 + \beta_1 x_i)}{1 + \exp(\beta_0 + \beta_1 x_i)} \right) = 0 \\ \frac{\partial L(\beta_0, \beta_1)}{\partial \beta_1} &= \sum_{i=1}^n x_i \left(y_i - \frac{\exp(\beta_0 + \beta_1 x_i)}{1 + \exp(\beta_0 + \beta_1 x_i)} \right) = 0 \end{aligned} \quad (3.11)$$

Uma vez que estas equações são não lineares é necessário recorrer a métodos numéricos para encontrar a solução, obtendo-se assim uma estimativa, $(\hat{\beta}_0, \hat{\beta}_1)$, de máxima verosimilhança

para β_0, β_1 . O método geralmente usado para resolver estas equações é o de *Newton-Raphson*; [10], [11].

3.1.2 Teste de Wald

O teste de Wald é utilizado para testar a significância estatística de um dado parâmetro. A estatística de teste utilizada que segue uma distribuição:

$$W = \frac{\hat{\beta}}{\sqrt{\text{var}(\hat{\beta})}} \sim N(0, 1), \quad (3.12)$$

onde $\sqrt{\text{var}(\hat{\beta})}$ representa o desvio padrão do estimador do parâmetro.

O valor W é depois comparado com valores tabulados de uma distribuição normal de acordo com o nível de significância pré-estabelecido, α .

3.1.3 Razão de chances (*Odds Ratio*)

Sabe-se que no modelo de regressão logística simples tem-se (3.7), pelo que:

$$\frac{\pi(x)}{1 - \pi(x)} = e^{\beta_0 + \beta_1 x}. \quad (3.13)$$

Variáveis Categóricas Independentes:

Esta medida é muito utilizada na área da saúde pois permite estudar se uma doença está ou não associada a um determinado fator. Para tal estudo, são necessários dois grupos (A e B) compostos por amostras independentes, onde o grupo A é composto por indivíduos com presença do fator ($x = 1$) e o grupo B é composto por indivíduos com ausência do fator ($x = 0$). Seja $\pi(x)$ a probabilidade da doença se desenvolver, a chance de desenvolvimento da doença para um indivíduo com a presença do fator é:

$$\frac{\pi(1)}{1 - \pi(1)} = e^{\beta_0 + \beta_1}, \quad (3.14)$$

por outro lado, a chance de desenvolvimento da doença para um indivíduo com a ausência do fator é:

$$\frac{\pi(0)}{1 - \pi(0)} = e^{\beta_0}. \quad (3.15)$$

Desta forma, a razão de chances de desenvolvimento da doença é dada por:

$$OR = \frac{\pi(1)\{1 - \pi(0)\}}{\pi(0)\{1 - \pi(1)\}} = e^{\beta_1}, \quad (3.16)$$

dependendo apenas do parâmetro β_1 .

Se estivermos na presença de uma variável preditora com mais de 2 categorias, ($k > 2$), para determinar OR utilizam-se $k - 1$ variáveis codificadoras conhecidas por variáveis

dummy. Normalmente, a primeira categoria é considerada como classe de referência e toma o valor zero para as $k - 1$ variáveis *dummy*. Nas variáveis *dummy* relativas às restantes categorias ($i = 2, \dots, k$) a entrada i toma o valor 1 e as restantes tomam o valor 0. Após este procedimento o cálculo do OR é efetuado de forma análoga.

Variáveis Contínuas Independentes:

No caso das variáveis contínuas, o estudo efetuado é relativo ao aumento de uma unidade na variável X_i , e não à presença ou ausência de um fator. Dessa forma a razão de chances será:

$$\begin{aligned}
 OR &= \frac{\text{Chance } (X_i \text{ aumenta uma unidade})}{\text{Chance } (X_i \text{ não varia})} \\
 &= \frac{\frac{P[Y=1|X_i=x+1]}{1-P[Y=1|X_i=x+1]}}{\frac{P[Y=1|X_i=x]}{1-P[Y=1|X_i=x]}} \\
 &= \frac{\exp(\beta_0 + \beta_1(X_i + 1))}{\exp(\beta_0 + \beta_1 X_i)} \\
 &= e^{\beta_1}
 \end{aligned} \tag{3.17}$$

Ou seja, quando X_i varia uma unidade, as chances de obter "sucesso" variam β_i unidades.

O intervalo de confiança para OR a 95% é dado por:

$$IC_{OR} = [\exp \beta_I, \exp \beta_S] \tag{3.18}$$

onde β_I e β_S são os limites inferior e superior, respetivamente, da estimativa de β .

Note-se ainda que:

- se $OR = 1$ não existem diferenças em ambos os grupos, isto é, o acontecimento é igualmente provável;
- se $OR > 1$ o evento é mais provável no 1º grupo;
- se $OR < 1$ o evento é menos provável no 1º grupo.

3.2 REGRESSÃO LOGÍSTICA MÚLTIPLA

Tal como acontece na regressão linear, também na regressão logística é possível considerar várias variáveis preditoras e não apenas só uma. Considerem-se então p variáveis preditoras X_1, X_2, \dots, X_p e seja $\mathbf{X} = (X_1, \dots, X_p)$. A probabilidade condicional da ocorrência de um evento é por definição $\pi(\mathbf{X})$ e a função de ligação é dada por:

$$\pi(\mathbf{X}) = \frac{\exp g(\mathbf{X})}{1 + \exp g(\mathbf{X})} \tag{3.19}$$

onde

$$g(\mathbf{X}) = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p. \tag{3.20}$$

3.2.1 Estimação dos Parâmetros

Para obter as estimativas dos parâmetros $\beta_i, i = 0, \dots, p$ do modelo utiliza-se o método de máxima verossimilhança, tal como visto na expressão (3.10) do modelo de regressão logística simples. De forma análoga e mais geral tem-se:

$$L(\beta) = \sum_{i=1}^n y_i(\beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi}) - \ln(1 + \exp(\beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi})). \quad (3.21)$$

Derivando a expressão 3.21, igualando a zero e substituindo pelos estimadores dos parâmetros obtém-se:

$$\begin{aligned} \frac{\partial L(\beta_0, \beta_1, \dots, \beta_p)}{\partial \beta_0} &= \sum_{i=1}^n \left(y_i - \frac{\exp(\beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi})}{1 + \exp(\beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi})} \right) = 0 \\ \frac{\partial L(\beta_0, \beta_1, \dots, \beta_p)}{\partial \beta_j} &= \sum_{i=1}^n x_j \left(y_i - \frac{\exp(\beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi})}{1 + \exp(\beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi})} \right) = 0, \forall j = 1, \dots, p \end{aligned} \quad (3.22)$$

Os estimadores dos parâmetros são dados pelas soluções destas $p+1$ equações, para as resolver utilizam-se métodos numéricos.

3.2.2 Teste de Wald

Assim como sucede na regressão logística simples, o teste de Wald utiliza-se para testar a significância estatística do valor de cada parâmetro populacional ser nulo. Desta forma, pretende-se testar:

$$H_0 : \beta_i = 0 \quad vs \quad H_1 : \beta_i \neq 0, \quad i = 0, \dots, p.$$

A estatística de teste é dada por:

$$W_i = \frac{\hat{\beta}_i}{\sqrt{\text{var}(\hat{\beta}_i)}}, \quad (3.23)$$

onde $\sqrt{\text{var}(\hat{\beta}_i)}$ representa o desvio padrão do estimador do parâmetro β_i .

O valor W é depois comparado com valores tabulados de uma normal padrão de acordo com o nível de significância pré-estabelecido, α .

3.2.3 Razão de chances ajustada (*Adjusted Odds Ratio*)

O procedimento para calcular a razão de chances no caso da regressão logística múltipla é similar ao utilizado no caso da regressão logística simples já exposto. A razão de chances ajustada (*adjusted odds ratio*), é obtida comparando indivíduos que diferem apenas numa característica de interesse e mantendo-se fixos os valores para todas as outras variáveis; [11].

3.2.4 Seleção de variáveis preditoras

Método Stepwise

Após a seleção das variáveis preditoras a serem incluídas no modelo de regressão logística múltiplo, importa saber quais são as mais importantes para explicar $\pi(x)$. A seleção de um modelo de regressão logística deve ser uma conjugação de seleção estatística de modelos e bom senso. Usar exclusivamente uma seleção estatística pode levar a que variáveis importantes para o estudo sejam excluídas do modelo por não terem significância estatística.

Existem vários algoritmos para a seleção das variáveis preditoras, mas na regressão logística o mais utilizado é o método *stepwise*. Este método tem como base um algoritmo de inclusão e/ou exclusão de variáveis preditoras segundo a importância das mesmas, tendo em consideração um certo critério estatístico. Por exemplo, o critério utilizado pode ser o teste de Wald para a significância de cada parâmetro β_i . Para cada um destes testes, deve ser imposto um valor de significância, P_S , pertencente ao intervalo $[0.15, 0.25]$ ¹, sendo que, quanto mais baixo for este valor de significância, mais importante será a variável preditora correspondente. Quando se incluem novas variáveis preditoras, outras que já foram incluídas podem ser dispensáveis, sendo necessário verificar a importância dessas, comparando os seus respectivos níveis com um limite superior P_S . Os coeficientes das variáveis preditoras com um nível de significância superior a P_S poderão ser candidatas à exclusão; [11].

Método Stepwise no RStudio

O método *stepwise* é também utilizado no *software RStudio* e existem três direções distintas para a seleção de variáveis preditoras: *backward*, *forward* e *both*. Os primeiros dois seguem direções opostas e o último é uma combinação dos dois. O método *forward* começa com o modelo mais simples, apenas com a constante β_0 , e vai construindo modelos cada vez mais complexos. Enquanto que o método *backward* começa com todas as variáveis preditoras e vai construindo modelos mais simples através da eliminação destas, chegando a um modelo onde já nenhuma pode ser eliminada segundo regras estabelecidas previamente.

3.2.5 Critérios de Informação

Os critérios abordados nesta secção serão o *Akaike Information Criterion* (AIC) e o *Bayesian Information Criterion* (BIC). Estes critérios constituem uma forma de avaliar um modelo estatístico.

Para seleccionar um modelo de entre os vários construídos, segundo o critério AIC, o escolhido deve ser o que possui menor valor. Caso existam dois modelos com o mesmo valor, deve ser escolhido o que tiver menor número de variáveis com poder preditivo.

Para os modelos de regressão múltipla, o AIC é dado por

$$AIC = -2\ln(l(\beta)) + 2p \quad (3.24)$$

onde l representa a função de verosimilhança e p é o número de variáveis do modelo.

¹Para mais esclarecimentos sobre o algoritmo utilizado, assim como o intervalo referido, consultar *Applied Logistic Regression*, Hosmer e Lemeshow.

Já o BIC é dado por:

$$BIC = -2 \ln(l(\beta)) + 2p \ln N \quad (3.25)$$

onde N é o número de observações. Também neste critério, deve ser escolhido o modelo com BIC mais baixo.

A diferença entre os dois critérios reside no termo de penalização, sendo este mais elevado no BIC. Este termo é igual a 2 no caso do AIC e igual a $\ln N$ no caso do BIC, ou seja, para o BIC o termo depende do número de observações; [15].

3.2.6 Teste de Hosmer & Lemeshow

O teste de Hosmer-Lemeshow foi proposto, pelos mesmos, como um teste à qualidade de ajustamento na regressão logística, tendo como hipótese nula que o modelo é o adequado e segundo Tattar (2017) [16] fornece resposta ao teste de hipóteses:

$$H_0 : E(Y) = \frac{\exp(\sum_{j=0}^p \beta_j x_{ij})}{1 + \exp(\sum_{j=0}^p \beta_j x_{ij})} \quad vs \quad H_1 : E(Y) \neq \frac{\exp(\sum_{j=0}^p \beta_j x_{ij})}{1 + \exp(\sum_{j=0}^p \beta_j x_{ij})}.$$

O teste de Hosmer-Lemeshow tem por base a divisão dos dados em g grupos de acordo com as probabilidades previstas e são propostas, pelos autores, duas formas de fazer esta divisão. A forma considerada ao longo do trabalho, consiste em agrupar com base nos percentis das probabilidades estimadas. Fixa-se $g = 10$ em que o primeiro grupo contém os $n'_1 = \frac{n}{10}$ indivíduos com as probabilidades de menor valor e o último grupo terá os $n'_{10} = \frac{n}{10}$ indivíduos com as probabilidades estimadas mais elevadas;[11].

A estatística de teste, \hat{C} , é dada por:

$$\hat{C} = \sum_{k=1}^g \frac{(O_k - n'_k \bar{\pi}_k)^2}{n'_k \bar{\pi}_k (1 - \bar{\pi}_k)} \quad (3.26)$$

onde $O_k = \sum_{j=1}^{C_k} Y_j$ com C_k o número de valores diferentes do conjunto das p variáveis preditoras observadas; n'_k é o número de indivíduos no k -ésimo grupo e $\bar{\pi}_k = \sum_{j=1}^{C_k} \frac{m_j \hat{\pi}_j}{n'_k}$, com m_j o número de indivíduos que possuem o mesmo conjunto de valores das variáveis preditoras, $x = x_j$ e $\hat{\pi}_j$ é a probabilidade estimada associada a j .

Segundo Hosmer e Lemeshow (2000) [11], a estatística \hat{C} segue uma distribuição aproximadamente de χ^2 com $g - 2$ graus de liberdade. O teste rejeita o ajuste do modelo quando $\hat{C}_{obs} > \chi^2_{g-2; 1-\alpha}$ para α pré-estabelecido.

3.2.7 Curva ROC

A análise ROC, *Receiver Operating Characteristic*, pode ser feita através de um gráfico que nos facilita o estudo da variação da sensibilidade, ou taxa de verdadeiros positivos, e 1-especificidade, ou taxa de verdadeiros negativos, para cada ponto de corte. A sensibilidade

é apresentada no eixo das ordenadas e 1-especificidade no eixo das abcissas. Para se encontrar o valor para o ponto de corte mais apropriado pode-se recorrer à curva ROC, que permite identificar o valor para o qual a sensibilidade e a especificidade do modelo se equilibram.

Uma forma de aferir a capacidade discriminante de um modelo a partir da curva ROC é através do cálculo da área representada abaixo da curva, *AUC*. Hosmer e Lemeshow (2000) [11] sugerem a seguinte classificação do poder discriminatório de um modelo de regressão logística:

- Se $AUC = 0.5$ o modelo não faz qualquer discriminação.
- Se $0.5 \leq AUC < 0.7$ o modelo apresenta uma discriminação fraca.
- Se $0.7 \leq AUC < 0.8$ o modelo apresenta uma discriminação aceitável.
- Se $0.8 \leq AUC < 0.9$ o modelo apresenta uma boa discriminação.
- Se $AUC \geq 0.9$ o modelo apresenta uma discriminação excecional.

Após a escolha do valor do ponto de corte, importa avaliar o poder de discriminação do modelo. Para tal foram considerados os conceitos: Precisão, Sensibilidade e Especificidade. A relação entre eles pode ser compreendida através da matriz de confusão:

Valor Estimado	Valor Observado	
	$Y = 0$	$Y = 1$
	$\hat{Y} = 0$	$\hat{Y} = 1$
	VN	FN
	FP	VP

Tabela 3.1: Matriz de confusão onde VN significa verdadeiros positivos, FN corresponde a falsos negativos, FP remete a falsos positivos e VP representa verdadeiros positivos

Tem-se:

-Precisão: Proporção de previsões corretas,

$$\text{Precisão} = \frac{VP + VN}{N} \quad (3.27)$$

onde N é o número total de observações.

-Sensibilidade: Proporção de verdadeiros positivos,

$$\text{Sensibilidade} = \frac{VP}{VP + FN}. \quad (3.28)$$

-Especificidade: Proporção de verdadeiros negativos,

$$\text{Especificidade} = \frac{VN}{VN + FP}. \quad (3.29)$$

Base de Dados

No decorrer do estágio foi fornecida uma base de dados proveniente de um certo hospital. Essa base de dados continha inicialmente 174 variáveis categóricas e numéricas. Dessas, foram eliminadas 141 porque apresentavam uma percentagem de valores omissos superior a 80%, com exceção da variável MORF_TUM. Esta variável não foi eliminada por se considerar importante para o estudo em causa. Ficando assim 33 variáveis para o estudo:

Nome da Variável	Descrição
AGE	Idade do utente
ADM_DIAG	Diagnóstico da Admissão
ADM_TIP	Tipo de Admissão
BIRTH_WGT	Peso à Nascimento
CASEMIX	Índice de case-mix
DDX1	Diagnóstico Principal (ICD9-MC)
DDX2, DDX3, DDX4 DDX5	Diagnósticos Secundários (ICD9-MC)
DEQ	Doentes Equivalentes
DSP	Código de Destino após alta
DSP_GDH	Código de Destino após alta conforme GDH
FIN_RESP	Entidade Financiadora
GCD_AP21	Código de Grande Categoria de Diagnóstico segundo o agrupador AP21
GCD_HCFA16	Código de Grande Categoria de Diagnóstico segundo o agrupador HCFA16
GDH_AP21	Código de GDH segundo o agrupador AP21 (grupo de diagnóstico homogéneo)
GDH_HCFA16	Código de GDH segundo o agrupador HCFA16 (grupo de diagnóstico homogéneo)
GDH_Tipo	Tipo de Grupo de diagnóstico homogéneo
ICU	Número de dias na unidade dos cuidados intensivos
MODULO	Módulo de atendimento
MORF_TUM	Código de Morfologia Tumoral
MOT_TRANF	Código de motivo de transferência
NUM_EPISOD	Número do Episódio
PRE_OP	Número de dias no pré-operatório
RINT	Reinternamento
SEXO	Sexo do indivíduo
SRG1, SRG2, SRG3, SRG4, SRG5	Procedimentos (ICD9-MC)
TOTDIAS	Total de dias de internamento

Tabela 4.1: Descrição das variáveis restantes após primeira análise

Optou-se por não usar nenhuma técnica de imputação de dados por se considerar não ser vantajoso para o estudo em questão. Considerando que a base de dados contém procedimentos, diagnósticos, entre outros registos de cada paciente, uma imputação poderia gerar dados de doentes irreais. Por exemplo, o doente ser do sexo masculino e ser codificado com um diagnóstico de complicações na gravidez.

De seguida apresenta-se uma análise detalhada sobre cada variável e de que forma se procedeu no sentido de melhorar a qualidade da base de dados.

1. Variável AGE: é quantitativa e corresponde à idade do doente.
2. Variável ADM_DIAG: é qualitativa e apresenta o código de diagnóstico principal existindo cerca de 15 000 códigos possíveis. Como existe um elevado número de códigos possíveis foi criada uma nova variável, DIA_ADM, que corresponde ao grupo de diagnóstico ao qual o código contido em ADM_DIAG pertence. Estes grupos foram criados conforme a secção de diagnóstico do documento ICD9CMOut2013¹. Posteriormente, a variável ADM_DIAG foi eliminada ficando a variável DIA_ADM representativa do diagnóstico de admissão. Esta nova variável, DIA_ADM, apresenta 18 categorias distintas:

Código	Descrição
1	Doenças Infecciosas e Parasitárias;
2	Neoplasias
3	Doenças Endócrinas, Nutricionais, Metabólicas e da Imunidade
4	Doenças do Sangue e dos Órgãos Hematopoiéticos
5	Doenças Mentais
6	Doenças do Sistema Nervoso e dos Órgãos dos Sentidos
7	Doenças do Aparelho Circulatório
8	Doenças do Aparelho Respiratório
9	Doenças do Aparelho Digestivo
10	Doenças do Aparelho Geniturinário
11	Complicações da Gravidez, do Parto e do Puerpério
12	Doenças da Pele e do Tecido Subcutâneo
13	Doenças do Aparelho Osteomuscular e do Tecido Conjuntivo
14	Anomalias Congénitas
15	Certas Condições originadas no Período Perinatal
16	Sintomas, Sinais e Condições Mal Definidas
17	Lesões e Intoxicações
18	Classificação Suplementar Factores com Influência no Estado de Saúde e Contacto com Serviços Saúde

Tabela 4.2: Codificação da variável DIA_ADM

¹Disponível no site: www2.acss.min-saude.pt/Portals/0/ICD9CMOut2013.xlsx

3. Variável ADM_TIP: é qualitativa e corresponde ao tipo de administração existindo sete códigos possíveis para o seu preenchimento:

Código	Descrição
1	Programada
2	Urgente
3	Acesso
4	Peclec
5	Medicina Privada
6	SIGIC
7	PACO

Tabela 4.3: Codificação da variável ADM_TIP

4. Variável BIRTH_WGT: é quantitativa e regista o peso à nascença.
5. Variável CASEMIX: é quantitativa e corresponde à parte que o doente representa para o índice de case-mix. Por ser uma variável que é preenchida com recurso ao GDH foi eliminada da base de dados.
6. Variável DDX1: é qualitativa e apresenta o código de diagnóstico principal existindo cerca de 15 000 códigos possíveis. Foram eliminadas as observações que não se encontravam preenchidas nesta variável, pois apenas seriam admitidos no estudo doentes com diagnóstico principal identificado. E à semelhança do que acontece com ADM_DIAG, foi criada uma nova variável, DIA_1, que corresponde ao grupo de diagnóstico ao qual o código contido em DDX1 pertence. A variável DDX1 foi eliminada ficando a variável DIA_1 representativa do diagnóstico principal. Esta nova variável, DIA_1, apresenta 18 categorias distintas tabeladas na Tabela 4.2.
7. Variáveis DDX2, DDX3, DDX4 e DDX5: são qualitativas e são responsáveis pelos diagnósticos secundários. À semelhança da variável DDX1 estas variáveis têm cerca de 15 000 códigos possíveis. Como tal, procedeu-se de forma idêntica: foram criadas quatro novas variáveis correspondentes a estas, DIA_2, DIA_3, DIA_4 e DIA_5, respetivamente, e as variáveis DDX2, DDX3, DDX4 e DDX5 foram eliminadas. Contudo, estes diagnósticos nem sempre são realizados. Como tal foi adicionada uma nova categoria, assim estas variáveis têm 19 categorias possíveis:

Código	Descrição
0	Diagnóstico não efetuado
1	Doenças Infecciosas e Parasitárias;
2	Neoplasias
3	Doenças Endócrinas, Nutricionais, Metabólicas e da Imunidade
4	Doenças do Sangue e dos Órgãos Hematopoiéticos
5	Doenças Mentais
6	Doenças do Sistema Nervoso e dos Órgãos dos Sentidos
7	Doenças do Aparelho Circulatório
8	Doenças do Aparelho Respiratório
9	Doenças do Aparelho Digestivo
10	Doenças do Aparelho Geniturinário
11	Complicações da Gravidez, do Parto e do Puerpério
12	Doenças da Pele e do Tecido Subcutâneo
13	Doenças do Aparelho Osteomuscular e do Tecido Conjuntivo
14	Anomalias Congénitas
15	Certas Condições originadas no Período Perinatal
16	Sintomas, Sinais e Condições Mal Definidas
17	Lesões e Intoxicações
18	Classificação Suplementar Factores com Influência no Estado de Saúde e Contacto com Serviços Saúde

Tabela 4.4: Codificação das variáveis DIA_2, DIA_3, DIA_4 e DIA_5

8. Variável DEQ: é quantitativa e corresponde à parte que o doente representa para o número de doentes equivalentes. Por ser uma variável que é preenchida com recurso ao GDH foi eliminada da base de dados.
9. Variável DSP: é qualitativa e apresenta o código de destino após a alta. Encontra-se codificada da seguinte forma:

Código	Descrição
1	Para o Domicílio
2	Para Outra Instituição com Internamento
6	Serviço Domiciliário
7	Saída Contra Parecer Médico
20	Falecido

Tabela 4.5: Codificação da variável DSP

Note-se que esta variável é registada no fim de um determinado episódio, do qual pode resultar o falecimento do doente. Fazendo então sentido existir o código 20.

10. Variável DSP_GDH: é qualitativa e representa o código de destino após alta conforme GDH. Por ter como base o GDH do doente foi também eliminada pois só é preenchida após a sua alta.

11. Variável FIN_RESP: é qualitativa e identifica qual a entidade responsável pelo financiamento. Foi eliminada, por sugestão da empresa, por não se considerar relevante para o estudo em questão.
12. Variáveis GCD_HCFA16 e GCD_AP21: são qualitativas e correspondem ao código da grande categoria de diagnóstico segundo o agrupador HCFA16 e ao código da grande categoria de diagnóstico segundo o agrupador AP21, respetivamente. Existem 25 grupos distintos e estas variáveis são registadas após o momento da alta, por essa razão, foram eliminadas da base de dados.
13. Variáveis GDH_HCFA16 e GDH_AP21: estas variáveis são ambas qualitativas e correspondem ao código do grupo de diagnóstico homogéneo segundo o agrupador HCFA16 e ao código do grupo de diagnóstico homogéneo segundo o agrupador AP21, respetivamente. Existem 25 grupos distintos e estas variáveis são registadas após o momento da alta, por essa razão, foram eliminadas da base de dados.
14. Variável GDH_TIP: é qualitativa que codifica o tipo de GDH: médico ou cirúrgico. Por ser uma variável que é preenchida com recurso ao GDH foi eliminada da base de dados.
15. Variável ICU: é quantitativa e regista o número de dias na unidade de cuidados intensivos.
16. Variável MODULO: é qualitativa e corresponde ao módulo onde foi atendido o doente. As observações que não correspondessem ao módulo de internamento foram eliminadas. Esta variável foi posteriormente eliminada.
17. Variável MORF_TUM: é qualitativa e apresenta o código da morfologia tumoral diagnosticada e existem mais de 200 categorias possíveis, sendo que apenas se encontra preenchida nos doentes que apresentam algum tumor. Então, esta foi recodificada da seguinte forma:

Tem tumor: 1

Não tem tumor: 0.

Assim, apenas avaliaremos de que forma a presença ou não de um tumor influencia o reinternamento.

18. Variável MOT_TRANF: é qualitativa e representa o código do motivo de transferência, existindo 5 códigos distintos onde um deles corresponde a não existir transferência.

Código	Descrição
0	Sem transferência
1	Realização de Exames
2	Para Seguimento
3	Por Falta de Recursos
4	Para Tratamento de condição associada

Tabela 4.6: Codificação da variável MOT_TRANF

19. Variável NUM_EPISOD: é qualitativa e corresponde ao número de episódio. Cada episódio é registado com um número e não existem dois episódios distintos com o mesmo número. Devido a isso, foram encontradas observações duplicadas que foram eliminadas. Esta variável foi posteriormente eliminada por ser meramente identificativa.

20. Variável PRE_OP: é quantitativa e regista o número de dias no pré-operatório.

21. Variável RINT: é qualitativa e representa se o episódio é um reinternamento ou não. Segue a codificação:

Código	Descrição
0	Não é reinternamento
1	É reinternamento

Tabela 4.7: Codificação da variável RINT

22. Variável SEXO: é qualitativa e representa o sexo de cada doente. É codificada da seguinte forma:

Código	Descrição
1	Masculino
2	Feminino

Tabela 4.8: Codificação da variável SEXO

Nesta variável existia ainda uma observação com a codificação 3 - Outro, esta observação foi eliminada por se tratar de um caso isolado.

23. Variáveis SRG1, SRG2, SRG3, SRG4 e SRG5: são qualitativas e correspondem aos procedimentos realizados. Seguem a mesma codificação, existindo cerca de 4 600 códigos possíveis. Uma vez que é um número bastante elevado de categorias, procedeu-se de forma semelhante à descrita anteriormente. Foram criadas cinco novas variáveis: PROC_1, PROC_2, PROC_3, PROC_4 e PROC_5 correspondendo, respetivamente, às variáveis SRG1, SRG2, SRG3, SRG4 e SRG5. Estas são preenchidas com o código correspondente ao grupo de procedimentos ao qual pertence o procedimento registado

nestas variáveis. Os grupos de procedimentos foram criados conforme a secção de diagnóstico do documento ICD9CMOut2013. As variáveis SRG1, SRG2, SRG3, SRG4 e SRG5 foram eliminadas ficando PROC_1, PROC_2, PROC_3, PROC_4 e PROC_5 representativas dos procedimentos efetuados, assumindo um dos seguintes códigos:

Código	Descrição
0	Procedimentos e intervenções não classificadas noutra parte
1	Procedimentos no sistema nervoso
2	Procedimentos no sistema endócrino
3	Procedimentos no olho
4	Outros procedimentos de diagnóstico e terapêutica não classificados noutra parte
5	Procedimentos nos ouvidos
6	Procedimentos no nariz, boca e faringe
7	Procedimentos no sistema respiratório
8	Procedimentos no sistema cardiovascular
9	Procedimentos no sistema linfático e imunológico
10	Procedimentos no sistema digestivo
11	Procedimentos do sistema urinário
12	Procedimentos nos órgãos genitais masculinas
13	Procedimentos nos órgãos genitais femininas
14	Procedimentos obstétricos
15	Procedimentos no sistema musculo-esquelético
16	Procedimentos no sistema integumentário
17	Outros procedimentos de diagnóstico e terapêutica não classificados
18	Procedimento não efetuado

Tabela 4.9: Codificação das variáveis PROC_1, PROC_2, PROC_3, PROC_4 e PROC_5

24. Variável TOTDIAS: é quantitativa e corresponde ao número total de dias de internamento.

Ficaram assim as seguintes 22 variáveis para estudo:

Nome da Variável	Descrição
AGE	Idade do utente
ADM_TIP	Tipo de Admissão
BIRTH_WGT	Peso à Nascimento
DIA_ADM	Grupo do Diagnóstico da Admissão
DIA_1	Grupo do Diagnóstico Principal
DIA_2, DIA_3, DIA_4 e DIA_5	Grupo do Diagnóstico Secundário
DSP	Código de Destino após alta
ICU	Número de dias na unidade dos cuidados intensivos
MORF_TUM	Código de Morfologia Tumoral
MOT_TRANF	Código de motivo de transferência
PRE_OP	Número de dias no pré-operatório
PROC_1, PROC_2, PROC_3, PROC_4 e PROC_5	Grupo do Procedimento
RINT	Reinternamento
SEXO	Sexo do indivíduo
TOTDIAS	Total de dias de internamento

Tabela 4.10: Variáveis finais

Assim, das 174 variáveis iniciais, ficam apenas 22 variáveis e dos 184 267 indivíduos iniciais restaram 24 782. Destes, verifica-se que 22 412 não são episódios de reinternamento e 2 370 são.

Resultados

Nesta secção será analisada a relação da variável RINT com as restantes da base de dados. Para isso, recorreu-se ao *software RStudio* e foram utilizadas as seguintes bibliotecas: `readr`, `lsr`, `ryouready`, `pROC`, `ResourceSelection`, `pscl`, `MASS`, `car` e `reshape2`. O principal objetivo passa por encontrar um modelo de previsão para o risco de um episódio ser de reinternamento que poderá ser consultado em qualquer momento. Desta forma não é necessário esperar pelos registos médicos do paciente em questão que por vezes se torna numa longa espera.

5.1 CORRELAÇÃO

Antes de compreender a relação entre as variáveis é necessário estudar a forma como elas se relacionam e se associam. Como referido anteriormente, na base de dados utilizada para este estudo, existem dois tipos de variáveis: qualitativas e quantitativas e por essa razão foram utilizadas as medidas de associação η^2 e Crámer. Estas medidas permitiram avaliar o nível de associação da variável de interesse RINT com as restantes da base de dados. (Tabela 5.1)

	RINT	Medida de associação
SEXO	0.0109	Crámer
MORF_TUM	0.0344	Crámer
DSP	0.1386	Crámer
BIRTH_WGT	0.0047	Eta
PRE_OP	0.0247	Eta
ICU	0.0145	Eta
ADM_TIP	0.0607	Crámer
TOTDIAS	0.0018	Eta
AGE	0.0433	Eta
MOT_TRANF	0.0179	Crámer
DIA_ADM	0.1196	Crámer
DIA_1	0.1244	Crámer
DIA_2	0.0866	Crámer
DIA_3	0.0748	Crámer
DIA_4	0.0676	Crámer
DIA_5	0.0747	Crámer
PROC_1	0.0944	Crámer
PROC_2	0.0549	Crámer
PROC_3	0.0413	Crámer
PROC_4	0.0425	Crámer
PROC_5	0.0457	Crámer

Tabela 5.1: Medidas de Associação

Da tabela acima constata-se não existirem valores que evidenciem fortes níveis de associação entre a variável RINT e as restantes, sendo que RINT se encontra mais associado com DSP, DIA_ADM e DIA_1. Observando valores de associação relativamente similares, optou-se por considerar todas as variáveis para a construção do modelo de regressão logística.

5.2 REGRESSÃO LOGÍSTICA SIMPLES

Inicialmente foi construído um modelo de regressão logística simples entre a variável RINT e cada uma das restantes variáveis preditoras, com o objetivo de avaliar a influência de cada variável no risco de se tratar de um reinternamento.

1. Variável Sexo:

	Não é reinternamento	Reinternamento	Total	Coefficiente β	Teste de Wald valor-p	OR IC a 95%
Masculino	11 834	1 296	13 130	Classe de Referência		
Feminino	10 578	1 074	11 652	-0.076	0.081	0.927 [0.852;1.009]
Total	22 412	2 370	24 782			

Tabela 5.2: Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável Sexo

Analisando a tabela verifica-se que dos 22 412 doentes que não se trata de reinternamento

11 834 são do sexo masculino e 10 578 são do sexo feminino e dos 2 370 que se trata de reinternamento, 1 296 são do sexo masculino e os restantes 1 074 são do sexo feminino.

Para a variável Sexo tem-se que o valor-p, relativo ao Teste de Wald, é inferior a 0.15 pelo que esta variável tem significância estatística ($\alpha = 0.15$) e será considerada no modelo de regressão logística múltiplo.

Note-se que não existe significância estatística, para $\alpha = 0.05$, relativamente à razão de chances para esta variável.

2. Variável MORF_TUM:

	Não é reinternamento	Reinternamento	Total	Coeficiente β	Teste de Wald valor-p	OR IC a 95%
0	21 560	2 332	23 892	Classe de Referência		
1	852	38	890	-0.886	1.17e-07	0.412 [0.297;0.572]
Total	22 412	2 370	24 782			

Tabela 5.3: Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável MORF_TUM

Através desta tabela verifica-se que dos 22 412 doentes que não são reinternamento 21 560 não tinham nenhum tipo de tumor e 852 tinham algum tipo de tumor e dos 2 370 reinternamentos, 2 332 não tinham nenhum tumor e os restantes 38 apresentavam esta doença.

Tem-se um valor-p inferior a 0.15, logo esta variável afeta significativamente o reinternamento. Assim será considerada no modelo de regressão logística múltiplo.

Com um valor de razão de chances de 0.412 e um respetivo intervalo a 95% confiança de [0.297;0.572], pode-se afirmar que os doentes com tumor têm uma chance de estarem num episódio de reinternamento 0.412 vezes menor à dos pacientes sem tumor. Ou seja, é menos provável um doente com tumor estar envolvido num episódio de reinternamento comparativamente aos doentes que não têm esta patologia.

3. Variável DSP:

	Não é reinternamento	Reinternamento	Total	Coeficiente β	Teste de Wald valor-p	OR IC a 95%
1	20 475	1 920	22 395	Classe de Referência		
2	439	60	499	0.377	0.007	1.458 [1.108;1.917]
6	450	32	482	-0.277	0.134	0.758 [0.528;1.089]
7	204	104	308	1.693	<2e-16	5.437 [4.273;6.916]
20	844	254	1 098	1.166	<2e-16	3.209 [2.768;3.721]
Total	22 412	2 370	24 782			

Tabela 5.4: Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável DSP

Observa-se que dos 22 412 doentes registados como sendo não é reinternamento, 20 475 foram para o domicílio após a alta, 439 seguiram para outra instituição com internamento, 450 serão acompanhados através de serviço domiciliário, 204 saíram contra parecer médico e 844 faleceram. Dos 2 370 que se verificou ser um reinternamento, 1 920 saíram para o domicílio, 60 foram para outra instituição com internamento, 32 serão acompanhados através de serviço domiciliário, 104 saíram contra parecer médico e 254 faleceram.

Uma vez que o valor-p obtido para o teste de Wald é inferior a 0.15 em pelo menos uma categoria, na verdade é inferior para todas as categorias, a variável é considerada relevante para o risco de ser reinternamento e será considerada no modelo múltiplo.

Destaca-se com um valor de razão de chances de 5.437 e um respetivo intervalo a 95% de confiança de [4.273; 6.916], que a chance de ser reinternamento é 5.437 maior nos pacientes que saem contra o parecer médico comparativamente à chance nos pacientes que seguem para o domicílio após a alta. Destaca-se também, com um valor de razão de chances de 3.209 e um respetivo intervalo a 95% de confiança de [2.768; 3.721], pelo que a chance de um dado episódio ser reinternamento é 3.209 vezes maior nos doentes que falecem nesse mesmo episódio comparativamente aos doentes que seguem para o seu domicílio.

4. Variável ADM_TIP:

	Não é reinternamento	Reinternamento	Total	Coeficiente β	Teste de Wald valor-p	OR IC a 95%
1	8 149	722	8 871	Classe de Referência		
2	13 593	1 635	15 228	0.306	6.65e-11	1.358 [1.239; 1.488]
6	670	13	683	-1.519	7.77e-08	0.219 [0.126; 0.381]
Total	22 412	2 370	24 782			

Tabela 5.5: Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável ADM_TIP

Note-se que apenas existem observações para as categorias 1, 2 e 6, por isso só é possível analisar os tipos de admissão: programada, urgente e admissão pelo SIGIC.

Observa-se que dos 22 412 doentes registados como sendo não é reinternamento, 8 149 tiveram a sua admissão programada, 13 593 foram admitidos em urgência e 670 foram admitidos pelo SIGIC. Dos 2 370 que se verificou ser um reinternamento, 722 tiveram a sua admissão programada, 1 635 foram admitidos em urgência e 13 foram admitidos pelo SIGIC.

Esta variável será incluída no modelo de regressão logística múltiplo por se encontrar dentro dos parâmetros estabelecidos, isto é, pelo menos uma categoria apresenta significância estatística já que o seu valor-p é inferior a 0.15.

Com um valor de razão de chances de 1.358 e um respetivo intervalo a 95% de confiança de [1.239; 1.488], pode-se afirmar que a chance de ser reinternamento é 1.358 vezes maior nos pacientes cuja admissão foi urgente do que os pacientes que têm a sua admissão programada.

5. Variável MOT_TRANF:

	Não é reinternamento	Reinternamento	Total	Coeficiente β	Teste de Wald valor-p	OR IC a 95%
0	21 957	2 309	24 266	Classe de Referência		
2	442	58	500	0.221	0.117	1.248 [0.946; 1.646]
3	6	0	6	-10.314	0.938	- -
4	7	3	10	1.405	0.042	4.075 [1.053; 15.777]
Total	22 412	2 370	24 782			

Tabela 5.6: Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável MOT_TRANF

Observa-se que dos 22 412 doentes registados como sendo não é reinternamento, 21 957 não foram transferidos, 442 foram transferidos para seguimento, 6 foram transferidos por falta de recursos e 7 foram transferidos para tratamento de condição associada. Dos 2 370 que se verificou ser um reinternamento, 2 309 não foram transferidos, 58 foram transferidos para seguimento e 3 foram transferidos para tratamento de condição associada. Note-se que não existem registos de transferências para realização de exames. Salienta-se ainda que não existem doentes que tenham sido transferidos por falta de recursos, categoria 3, e se trate de

um episódio de reinternamento.

Como o valor-p do teste de Wald é inferior a 0.15 nas categorias 2 e 4, a variável será utilizada para a obtenção de um modelo de regressão logística múltiplo.

Com um valor de razão de chances de 4.075 e um respectivo intervalo a 95% de confiança de [1.053; 15.777], pode-se afirmar que a chance de ser reinternamento é 4.075 vezes maior nos doentes que são transferidos para tratamento de doença associada relativamente à chance nos que não são transferidos.

6. Variável DIA_ADM:

	Não é reinternamento	Reinternamento	Total	Coeficiente β	Teste Wald valor-p	OR IC a 95%
1	28	1	29	-1.578	0.120827	0.206 [0.028; 1.516]
2	2 596	450	3 046	Classe de Referência		
3	732	69	801	-0.609	7.36e-06	0.544 [0.417; 0.710]
4	175	23	198	-0.277	0.224	0.758 [0.485; 1.184]
5	340	21	361	-1.032	7.62e-06	0.356 [0.227; 0.560]
6	1 258	65	1 323	-1.210	<2e-16	0.298 [0.228; 0.390]
7	3 515	267	3 782	-0.825	<2e-16	0.438 [0.374; 0.514]
8	2 827	297	3 124	-0.501	3.08e-10	0.606 [0.519; 0.708]
9	2 276	278	2 554	-0.350	1.75e-05	0.705 [0.601; 0.827]
10	2 230	173	2 403	-0.804	<2e-16	0.448 [0.372; 0.538]
11	447	68	515	-0.131	0.350	0.878 [0.667; 1.154]
12	219	17	236	-0.803	0.002	0.448 [0.271; 0.741]
13	571	40	611	-0.906	1.24e-07	0.404 [0.289; 0.565]
14	446	28	474	-1.016	4.59e-07	0.362 [0.244; 0.537]
15	32	33	65	1.783	1.92e-12	5.949 [3.621; 9.774]
16	1 405	173	1 578	-0.342	0.000336	0.710 [0.589; 0.856]
17	1 608	215	1 823	-0.260	0.003	0.771 [0.648; 0.918]
18	1 707	152	1 859	-0.666	1.60e-11	0.514 [0.423; 0.624]
Total	22 412	2 370	24 782			

Tabela 5.7: Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável DIA_ADM

Como todas as categorias da variável são relativas a códigos de diagnósticos efetuados, optou-se

por usar a categoria 2, neoplasias, como classe de referência. Esta escolha resulta da realização de vários modelos de regressão logística simples usando diferentes categorias como classe de referência e foi com a categoria 2 que se obteve maior número de significâncias estatísticas nas restantes categorias.

Através da análise da tabela é possível observar a distribuição dos doentes pelos vários códigos de diagnóstico.

Como existe pelo menos uma categoria com relevância estatística, isto é com valor-p inferior a 0.15, esta variável será utilizada para a construção do modelo de regressão logística múltiplo. Sobressai o valor de razão de chances de 5.949 e o respetivo intervalo a 95% confiança de [3.621; 9.774], pelo que a chance de ser reinternamento é cerca de 6 vezes maior nos doentes que têm certas condições originadas no período perinatal do que a chance nos doentes que sofrem de neoplasias. Note-se que comparativamente às restantes categorias o risco de ser reinternamento é mais provável para os doentes que sofrem de neoplasias já que o valor de razão de chances é inferior a 1. Porém, as categorias 1, 4 e 11 não são significativas ao nível de significância de 5%.

7. Variável DIA_1

	Não é reinternamento	Reinternamento	Total	Coeficiente β	Teste Wald valor-p	OR IC a 95%
1	49	4	53	-0.657	0.208	0.518 [0.186; 1.443]
2	3 226	508	3 734	Classe de Referência		
3	795	73	868	-0.534	3.98e-05	0.583 [0.451; 0.754]
4	144	19	163	-0.177	0.477	0.838 [0.515; 1.364]
5	367	21	388	-1.012	1.02e-05	0.363 [0.232; 0.570]
6	1 340	81	1 421	-0.957	1.14e-14	0.384 [0.301; 0.489]
7	4 173	301	4 474	-0.781	<2e-16	0.458 [0.394; 0.532]
8	2 837	282	3 119	-0.460	4.80e-09	0.631 [0.541; 0.736]
9	2 313	279	2 592	-0.267	0.00078	0.766 [0.656; 0.895]
10	1 781	126	1 907	-0.800	1.28e-14	0.449 [0.367; 0.551]
11	633	84	717	-0.171	0.173	0.843 [0.659; 1.078]
12	193	15	208	-0.706	0.010	0.494 [0.289; 0.842]
13	614	37	651	-0.961	4.72e-08	0.383 [0.271; 0.540]
14	546	30	576	-1.053	5.28e-08	0.349 [0.239; 0.510]
15	29	33	62	1.978	2.22e-14	7.226 [4.350; 12.005]
16	273	30	303	-0.360	0.069	0.698 [0.473; 1.029]
17	1 966	311	2 277	0.005	0.953	1.005 [0.863; 1.169]
18	1 133	136	1 269	-0.271	0.008	0.762 [0.623; 0.932]
Total	22 412	2 370	24 782			

Tabela 5.8: Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável DIA_1

Tal como anteriormente, como todas as categorias da variável são relativas a diagnósticos efetuados, optou-se por usar a categoria 2, neoplasias, como classe de referência. Esta escolha resulta da realização de vários modelos de regressão logística simples usando diferentes categorias como classe de referência e foi com a categoria 2 que se obteve maior número de significâncias estatísticas nas restantes categorias.

Através da análise da tabela é possível observar a distribuição dos doentes pelos vários códigos de diagnóstico.

Como existe pelo menos uma categoria com relevância estatística, isto é com valor-p inferior a 0.15, esta variável será utilizada para a construção do modelo de regressão logística múltiplo.

Destaca-se o valor de razão de chances de 7.226 e o respectivo intervalo a 95% confiança de [4.350; 12.005]; pode-se afirmar que a chance de ser reinternamento é 7.226 vezes maior nos doentes que têm certas condições originadas no período perinatal relativamente à chance nos doentes que sofrem de neoplasias. Observe-se que as categorias 1, 4, 11, 16 e 17 não são significativas ao nível de significância de 5% já que o 1 está contido nos seus intervalos a 95% de confiança.

8. Variável DIA_2

	Não é reinternamento	Reinternamento	Total	Coeficiente β	Teste Wald valor-p	OR IC a 95%
0	1 645	145	1 790	Classe de Referência		
1	102	11	113	0.202	0.540	1.223 [0.642; 2.331]
2	451	101	552	0.932	2.81e-11	2.541 [1.931; 3.343]
3	1 763	132	1 895	-0.163	0.192	0.849 [0.665; 1.085]
4	575	106	681	0.738	6.71e-08	2.091 [1.600; 2.734]
5	877	81	958	0.047	0.747	1.048 [0.789; 1.392]
6	813	92	905	0.250	0.074	1.284 [0.976; 1.689]
7	2 971	256	3 227	-0.023	0.834	0.978 [0.790; 1.209]
8	760	73	833	0.086	0.567	1.090 [0.812; 1.462]
9	1 042	121	1 163	0.276	0.033	1.317 [1.022; 1.697]
10	1 206	104	1 310	-0.022	0.870	0.978 [0.752; 1.272]
11	154	18	172	0.282	0.285	1.326 [0.791; 2.223]
12	327	38	365	0.276	0.150	1.318 [0.905; 1.921]
13	480	40	520	-0.056	0.763	0.945 [0.657; 1.361]
14	309	26	335	-0.046	0.834	0.955 [0.618; 1.474]
15	217	9	226	-0.754	0.032	0.471 [0.237; 0.936]
16	1 801	305	2 106	0.653	8.66e-10	1.921 [1.559; 2.367]
17	487	57	544	0.284	0.085	1.328 [0.962; 1.833]
18	6 432	655	7 087	0.144	0.132	1.155 [0.957; 1.394]
Total	22 412	2 370	24 782			

Tabela 5.9: Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável DIA_2

Analisando a tabela é possível observar a distribuição dos doentes pelos vários códigos de diagnóstico. Nesta variável optou-se por usar como classe de referência a categoria 0 que representa diagnóstico não efetuado. Assim, é possível analisar a diferença entre ter diagnóstico secundário efetuado e não efetuado.

Como existe pelo menos uma categoria com relevância estatística esta variável será utilizada para a construção do modelo de regressão logística múltiplo.

Com um valor de razão de chances de 2.541 e um respetivo intervalo a 95% confiança de [1.931; 3.343], pode-se afirmar que a chance de ser reinternamento é 2.541 vezes maior nos doentes que têm neoplasias do que a chance nos doentes cujo diagnóstico não foi efetuado. As categorias 4, 9 e 16 apresentam um valor de chances superior a 1, assim o risco de ser reinternamento é maior para os doentes com o diagnóstico correspondente às categorias referidas do que os doentes cujo diagnóstico não foi efetuado. Já as restantes categorias não são significativas ao nível de significância de 5%.

9. Variável DIA_3

	Não é reinternamento	Reinternamento	Total	Coeficiente β	Teste Wald valor-p	OR IC a 95%
0	3 765	355	4 120	Classe de Referência		
1	103	20	123	0.722	0.004	2.059 [1.260; 3.365]
2	384	76	460	0.741	6.61e-08	2.099 [1.604; 2.747]
3	1 545	139	1 684	-0.047	0.653	0.954 [0.777; 1.171]
4	618	104	722	0.579	1.29e-06	1.785 [1.412; 2.256]
5	818	57	875	-0.302	0.041	0.739 [0.553; 0.987]
6	849	81	930	0.012	0.927	1.012 [0.786; 1.303]
7	3 194	264	3 458	-0.132	0.120	0.877 [0.742; 1.035]
8	883	87	970	0.044	0.726	1.045 [0.817; 1.336]
9	998	137	1 135	0.376	0.0004	1.456 [1.181; 1.795]
10	1 049	108	1 157	0.088	0.446	1.092 [0.871; 1.369]
11	177	22	199	0.276	0.235	1.318 [0.835; 2.080]
12	373	55	428	0.447	0.004	1.564 [1.155; 2.118]
13	405	34	439	-0.116	0.535	0.890 [0.617; 1.284]
14	255	18	273	-0.290	0.247	0.749 [0.459; 1.222]
15	115	5	120	-0.774	0.092	0.461 [0.187; 1.136]
16	1 889	280	2 169	0.452	9.43e-08	1.572 [1.331; 1.856]
17	500	66	566	0.336	0.018	1.400 [1.059; 1.850]
18	4 492	462	4 954	0.087	0.240	1.091 [0.944; 1.261]
Total	22 412	2 370	24 782			

Tabela 5.10: Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável DIA_3

À semelhança do explicado na variável DIA_2, também para esta variável se optou por usar a categoria 0 como classe de referência. Através da tabela é também possível observar a distribuição dos doentes pelos vários grupos de diagnóstico.

Uma vez que existe pelo menos uma categoria com relevância estatística, esta variável será utilizada para a construção do modelo de regressão logística múltiplo.

Evidencia-se o valor de razão de chances de 2.099 e o respetivo intervalo a 95% confiança de [1.604; 2.747], pelo que a chance de ser reinternamento é 2.099 vezes nos doentes que

têm neoplasias à chance nos doentes cujo diagnóstico não foi efetuado. O mesmo se verifica também, embora com valores distintos, para as categorias 1, 4, 9, 12, 16 e 17. As restantes categorias, com exceção da categoria 5, não são significativas ao nível de significância de 5%.

10. Variável DIA_4

	Não é reinternamento	Reinternamento	Total	Coeficiente β	Teste Wald valor-p	OR IC a 95%
0	5 872	550	6 422	Classe de Referência		
1	87	13	100	0.467	0.120	1.595 [0.885; 2.876]
2	348	67	415	0.721	3.02e-07	2.056 [1.560; 2.708]
3	1 531	135	1 666	-0.060	0.547	0.941 [0.773; 1.146]
4	557	100	657	0.651	2.99e-08	1.917 [1.523; 2.413]
5	667	56	723	-0.109	0.454	0.896 [0.673; 1.194]
6	725	70	795	0.030	0.819	1.031 [0.794; 1.337]
7	3 264	285	3 549	-0.070	0.357	0.932 [0.803; 1.082]
8	800	90	890	0.183	0.126	1.201 [0.950; 1.519]
9	880	112	992	0.307	0.005	1.359 [1.096; 1.685]
10	915	90	1 005	0.049	0.681	1.050 [0.831; 1.326]
11	121	13	134	0.137	0.642	1.147 [0.643; 2.046]
12	323	38	361	0.228	0.198	1.256 [0.888; 1.778]
13	314	27	341	-0.086	0.677	0.918 [0.614; 1.373]
14	185	12	197	-0.367	0.223	0.693 [0.384; 1.250]
15	73	5	78	-0.313	0.500	0.731 [0.294; 1.817]
16	1 789	268	2 057	0.470	3.10e-09	1.599 [1.369; 1.868]
17	376	54	430	0.427	0.005	1.533 [1.138; 2.066]
18	3 585	385	3 970	0.137	0.050	1.147 [1.000; 1.315]
Total	22 412	2 370	24 782			

Tabela 5.11: Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável DIA_4

Tal como anteriormente, é possível analisar a distribuição dos doentes pelos vários códigos de diagnóstico, sendo que também para esta variável foi utilizada a categoria 0 como classe de referência.

Como existe pelo menos uma categoria com relevância estatística, isto é com valor-p inferior a 0.15, esta variável será utilizada para a construção do modelo de regressão logística múltiplo. Distingue-se o valor de razão de chances de 2.056 e o respetivo intervalo a 95% confiança de [1.560; 2.708], pelo que a chance de ser reinternamento é 2.056 vezes maior nos doentes que têm neoplasias do que a chance nos doentes cujo diagnóstico não foi efetuado. O mesmo também se verifica, embora com valores distintos, para as categorias 4, 9, 16, 17 e 18. As restantes não são significativas ao nível de significância de 5%.

11. Variável DIA_5

	Não é reinternamento	Reinternamento	Total	Coeficiente β	Teste Wald valor-p	OR IC a 95%
0	7 723	740	8 463	Classe de Referência		
1	79	11	90	0.374	0.249	1.453 [0.770; 2.743]
2	317	57	374	0.629	2.37e-05	1.877 [1.402; 2.513]
3	1 508	131	1 639	-0.098	0.321	0.907 [0.747; 1.101]
4	564	116	680	0.764	2.39e-12	2.147 [1.734; 2.658]
5	573	68	641	0.214	0.110	1.239 [0.953; 1.610]
6	705	66	771	-0.023	0.863	0.977 [0.751; 1.271]
7	3 126	275	3 401	-0.085	0.247	0.918 [0.795; 1.061]
8	753	87	840	0.187	0.118	1.206 [0.954; 1.524]
9	773	109	882	0.386	0.0004	1.472 [1.188; 1.823]
10	801	90	891	0.159	0.176	1.173 [0.931; 1.477]
11	84	12	96	0.399	0.199	1.491 [0.810; 2.743]
12	260	37	297	0.396	0.028	1.485 [1.044; 2.113]
13	300	22	322	-0.267	0.233	0.765 [0.493; 1.188]
14	163	11	174	-0.351	0.264	0.704 [0.381; 1.303]
15	50	2	52	-0.874	0.226	0.417 [0.101; 1.718]
16	1 521	228	1 749	0.448	3.01e-08	1.564 [1.335; 1.833]
17	292	49	341	0.560	0.0004	1.751 [1.282; 2.392]
18	2 820	259	3 079	-0.042	0.575	0.959 [0.827; 1.111]
Total	22 412	2370	24 782			

Tabela 5.12: Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável DIA_5

Analisando a tabela, observa-se a distribuição dos doentes pelos vários códigos de diagnóstico. Note-se que também para a variável DIA_5 usou-se, como classe de referência, a categoria 0. Já que existe pelo menos uma categoria com relevância estatística, isto é, com valor-p inferior a 0.15, esta variável será utilizada para a construção do modelo de regressão logística múltiplo. Destaca-se o valor de razão de chances de 2.147 e o respetivo intervalo a 95% confiança de [1.734; 2.658], pelo que a chance de ser reinternamento é 2.147 vezes maior nos doentes que têm doenças no sangue e dos órgãos hematopoiéticos do que a chance nos doentes cujo diagnóstico não foi efetuado. O mesmo também se verifica, embora com valores distintos, para as categorias 2, 9, 12, 16 e 17. As restantes categorias não são significativas ao nível de significância de 5% já que o 1 está contido nos seus intervalos a 95% de confiança.

12. Variável PROC_1

	Não é reinternamento	Reinternamento	Total	Coeficiente β	Teste Wald valor-p	OR IC a 95%
3	816	27	843	-2.120	8.55e-15	0.120 [0.070; 0.205]
5	173	21	194	-0.820	0.006	0.440 [0.245; 0.793]
6	349	23	372	-1.431	6.61e-07	0.239 [0.136; 0.420]
7	695	65	760	-1.081	2.84e-06	0.339 [0.216; 0.533]
8	2 190	154	2 344	-1.366	5.50e-11	0.255 [0.170; 0.384]
9	194	23	217	-0.844	0.004	0.430 [0.243; 0.762]
10	2 263	288	2 551	-0.773	0.000120	0.462 [0.311; 0.685]
11	1 128	75	1 203	-1.422	2.67e-10	0.241 [0.155; 0.375]
12	357	14	371	-1.950	4.56e-09	0.142 [0.074; 0.273]
13	304	31	335	-0.994	0.000211	0.370 [0.219; 0.626]
14	398	46	444	-0.869	0.000420	0.419 [0.259; 0.680]
15	1 306	71	1 377	-1.623	7.67e-13	0.197 [0.127; 0.307]
16	667	43	710	-1.453	4.30e-09	0.234 [0.144; 0.380]
17	11 445	1 454	12 899	-0.774	5.97e-05	0.461 [0.316; 0.673]
18	127	35	162	Classe de Referência		
Total	22 412	2 370	24 782			

Tabela 5.13: Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável PROC_1

Através da tabela é possível analisar a distribuição dos doentes pelos vários códigos de procedimentos. Note-se que não existem doentes registados nas categorias 0 - Procedimentos

e intervenções não classificadas noutra parte, 1 - Procedimentos no sistema nervoso, 2 - Procedimentos no sistema endócrino e 4 - Outros procedimentos de diagnóstico e terapêutica não classificados noutra parte.

Esta variável será utilizada para a construção do modelo de regressão logística múltiplo, porque existe pelo menos uma categoria com relevância estatística.

Com 95% de confiança, e com valores de razão de chances todos inferiores a 1, pode-se afirmar que independentemente do procedimento realizado, o risco de ser reinternamento é inferior para os doentes que são submetidos a qualquer procedimento do que para os doentes que não são submetidos a nenhum procedimento.

13. Variável PROC_2

	Não é reinternamento	Reinternamento	Total	Coeficiente β	Teste Wald valor-p	OR IC a 95%
3	500	11	511	-1.451	3.51e-06	0.234 [0.127; 0.433]
5	171	12	183	-0.291	0.342	0.747 [0.409; 1.364]
6	172	21	193	0.262	0.278	1.300 [0.810; 2.088]
7	330	39	369	0.230	0.210	1.258 [0.878; 1.803]
8	1 097	106	1 203	0.029	0.818	1.029 [0.807; 1.312]
9	193	17	210	-0.064	0.807	0.938 [0.561; 1.570]
10	825	93	918	0.183	0.161	1.200 [0.930; 1.549]
11	679	44	723	-0.371	0.030	0.690 [0.494; 0.965]
12	55	3	58	-0.543	0.363	0.581 [0.180; 1.872]
13	163	16	179	0.044	0.870	1.045 [0.614; 1.779]
14	246	26	272	0.118	0.588	1.125 [0.734; 1.725]
15	411	26	437	-0.395	0.065	0.674 [0.443; 1.025]
16	315	19	334	-0.443	0.073	0.642 [0.396; 1.041]
17	14 923	1 718	16 641	0.204	0.007	1.226 [1.058; 1.420]
18	2 332	219	2551	Classe de Referência		
Total	22 412	2 370	24 782			

Tabela 5.14: Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável PROC_2

Tal como anteriormente, é possível observar e analisar a distribuição dos doentes pelos vários códigos de procedimentos. Também aqui não existem doentes registados nas categorias 0 - Procedimentos e intervenções não classificadas noutra parte, 1 - Procedimentos no sistema

nervoso, 2 - Procedimentos no sistema endócrino e 4 - Outros procedimentos de diagnóstico e terapêutica não classificados noutra parte.

Como existe pelo menos uma categoria com relevância estatística, isto é com valor-p inferior a 0.15, esta variável será utilizada para a construção do modelo de regressão logística múltiplo. Com um valor de razão de chances 1.226 e o intervalo a 95% de confiança [1.058; 1.420], pode-se afirmar que a chance de ser reinternamento é maior 1.226 vezes nos doentes que foram submetidos a um procedimento do grupo outros procedimentos de diagnóstico e terapêutica não classificados comparativamente à chance dos doentes que não foram submetidos a um procedimento. As restantes categorias, com exceção das 3 e 11, não são significativas ao nível de significância de 5%.

14. Variável PROC_3

	Não é reinternamento	Reinternamento	Total	Coeficiente β	Teste Wald valor-p	OR IC a 95%
3	311	10	321	-1.089	0.000803	0.336 [0.178; 0.636]
5	118	12	130	0.062	0.840	1.064 [0.583; 1.943]
6	157	12	169	-0.224	0.462	0.800 [0.441; 1.451]
7	203	25	228	0.253	0.245	1.288 [0.840; 1.976]
8	995	84	1 079	-0.124	0.320	0.883 [0.692; 1.128]
9	117	12	129	0.071	0.819	1.073 [0.587; 1.960]
10	479	54	533	0.165	0.279	1.179 [0.875; 1.590]
11	444	34	478	-0.222	0.231	0.801 [0.557; 1.152]
12	19	1	20	-0.597	0.561	0.551 [0.074; 4.123]
13	92	6	98	-0.382	0.368	0.682 [0.297; 1.568]
14	114	9	123	-0.191	0.585	0.826 [0.416; 1.640]
15	206	13	219	-0.415	0.153	0.660 [0.374; 1.167]
16	169	17	186	0.051	0.844	1.052 [0.633; 1.751]
17	14 667	1 668	16 335	0.174	0.003	1.190 [1.063; 1.332]
18	4 321	413	4 734	Classe de Referência		
Total	22 412	2 370	24 782			

Tabela 5.15: Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável PROC_3

Analisando a tabela é possível entender a distribuição dos doentes pelos vários códigos de procedimentos. Tal como anteriormente, para esta variável não existem doentes registados nas

categorias 0 - Procedimentos e intervenções não classificadas noutra parte, 1 - Procedimentos no sistema nervoso, 2 - Procedimentos no sistema endócrino e 4 - Outros procedimentos de diagnóstico e terapêutica não classificados noutra parte.

Já que existe pelo menos uma categoria com relevância estatística, isto é com valor-p inferior a 0.15, esta variável será utilizada para a construção do modelo de regressão logística múltiplo. Realça-se que apenas duas categorias, 3 e 17, são significativas ao nível de significância de 5%, destacando-se o valor de razão de chances de 1.288 e o respetivo intervalo a 95% de confiança de [0.840; 1.976], pelo que a chance de ser reinternamento é 1.288 vezes maior nos doentes que são submetidos a procedimentos no sistema respiratório do que os doentes que não foram submetidos a um procedimento.

15. Variável PROC_4

	Não é reinternamento	Reinternamento	Total	Coeficiente β	Teste Wald valor-p	OR IC a 95%
3	196	8	204	-0.841	0.021	0.431 [0.212; 0.879]
5	91	10	101	0.150	0.656	1.161 [0.601; 2.243]
6	66	3	69	-0.733	0.215	0.480 [0.151; 1.532]
7	172	21	193	0.255	0.278	1.290 [0.814; 2.045]
8	796	72	868	-0.045	0.730	0.956 [0.741; 1.234]
9	132	6	138	-0.733	0.081	0.480 [0.211; 1.093]
10	348	38	386	0.143	0.416	1.154 [0.817; 1.630]
11	354	28	382	-0.179	0.373	0.836 [0.564; 1.239]
12	15	1	16	-0.350	0.735	0.705 [0.093; 5.343]
13	39	1	40	-1.306	0.196	0.271 [0.037; 1.963]
14	88	9	97	0.078	0.825	1.081 [0.542; 2.157]
15	129	6	135	-0.710	0.091	0.492 [0.216; 1.119]
16	108	6	114	-0.532	0.207	0.587 [0.257; 1.342]
17	13 473	1 555	15 028	0.199	7.61e-05	1.220 [1.105; 1.346]
18	6 405	606	7 011	Classe de Referência		
Total	22 412	2 370	24 782			

Tabela 5.16: Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável PROC_4

Através da tabela, é possível observar e analisar a distribuição dos doentes pelos vários códigos de procedimentos. Note-se que também não existem doentes registados nas categorias 0 -

Procedimentos e intervenções não classificadas noutra parte, 1 - Procedimentos no sistema nervoso, 2 - Procedimentos no sistema endócrino e 4 - Outros procedimentos de diagnóstico e terapêutica não classificados noutra parte.

Como existe pelo menos uma categoria com relevância estatística, isto é com valor-p inferior a 0.15, esta variável será utilizada para a construção do modelo de regressão logística múltiplo. Aponta-se que apenas duas categorias, 3 e 17, são significativas ao nível de significância de 5%, destacando-se o valor de razão de chances de 1.220 e o respetivo intervalo a 95% de confiança de [1.105; 1.346], pelo que a chance de ser reinternamento é 1.220 vezes maior nos doentes que são submetidos a procedimentos no sistema respiratório do que os doentes que não foram submetidos a um procedimento.

16. Variável PROC_5

	Não é reinternamento	Reinternamento	Total	Coeficiente β	Teste Wald valor-p	OR IC a 95%
3	102	6	108	-0.468	0.267	0.626 [0.274; 1.431]
5	51	9	60	0.630	0.083	1.878 [0.921; 3.829]
6	45	5	50	0.168	0.723	1.183 [0.468; 2.988]
7	153	22	175	0.426	0.066	1.530 [0.973; 2.407]
8	569	45	614	-0.172	0.279	0.842 [0.616; 1.150]
9	66	2	68	-1.132	0.115	0.323 [0.079; 1.319]
10	278	27	305	0.033	0.872	1.034 [0.692; 1.545]
11	358	20	378	-0.520	0.026	0.595 [0.377; 0.938]
12	5	0	5	-10.201	0.944	- -
13	6	0	6	-10.201	0.939	- -
14	97	5	102	-0.600	0.192	0.549 [0.223; 1.352]
15	65	6	71	-0.018	0.967	0.982 [0.424; 2.274]
16	61	3	64	-0.647	0.275	0.523 [0.164; 1.672]
17	12 084	1 424	13 508	0.227	1.09e-06	1.254 [1.145; 1.374]
18	8 472	796	9 268	Classe de Referência		
Total	22 412	2 370	24 782			

Tabela 5.17: Tabela de contingência e estimativas associadas ao modelo de regressão logística simples tomando a variável PROC_5

Tal como anteriormente, é possível observar a distribuição dos doentes pelos vários códigos de procedimentos. Também para esta variável não existem doentes registados nas categorias 0 -

Procedimentos e intervenções não classificadas noutra parte, 1 - Procedimentos no sistema nervoso, 2 - Procedimentos no sistema endócrino e 4 - Outros procedimentos de diagnóstico e terapêutica não classificados noutra parte. Observa-se ainda que não existem indivíduos que estejam num episódio de reinternamento e tenham sido submetidos a um procedimento dos grupos 12 - Procedimentos nos órgãos genitais masculinas e 13 - Procedimentos nos órgãos genitais femininas.

Como existe pelo menos uma categoria com relevância estatística, isto é com valor-p inferior a 0.15, esta variável será utilizada para a construção do modelo de regressão logística múltiplo. Aponta-se que apenas duas categorias, 11 e 17, são significativas ao nível de significância de 5%, destacando-se o valor de razão de chances de 1.254 e o respetivo intervalo a 95% de confiança de [1.145; 1.374], pelo que a chance de ser reinternamento é 1.254 vezes maior nos doentes que são submetidos a procedimentos no sistema respiratório do que os doentes que não foram submetidos a um procedimento.

17. Variáveis Numéricas:

	Coefficiente β	Teste de Wald valor-p	OR IC a 95%
BIRTH_WGT	-3.555e-05	0.459	1.000 [0.9999; 1.0001]
PRE_OP	-0.028365	9.54e-05	0.109 [0.104; 0.114]
ICU	-0.03138	0.0164	0.969 [0.945; 0.994]
TOTDIAS	-0.0004492	0.776	1.000 [0.996; 1.003]
AGE	-0.0055333	1.07e-11	0.994 [0.993; 0.996]

Tabela 5.18: Estimativas associadas aos modelos de regressão logística simples tomando cada uma das variáveis numéricas

Das variáveis numéricas apenas serão consideradas para o modelo de regressão logísticas as seguintes: PRE_OP, ICU e AGE. As restantes serão descartadas já que têm os respetivos valores-p superiores a 0.15, o que evidencia não existir efeito estatisticamente significativo destas variáveis sobre o risco de ser um episódio de reinternamento.

Com um valor de razão de chances 0.969 e um respetivo intervalo a 95% de confiança [0.945; 0.994], pode-se afirmar que adicionando um dia na unidade de cuidados intensivos a chance de ser reinternamento é 0.969 vezes menor.

5.3 REGRESSÃO LOGÍSTICA MÚLTIPLA

Com o intuito de obter o melhor modelo de regressão logística múltiplo possível para a previsão do risco de ser reinternamento, foram construídos 12 modelos distintos. O procedimento encontra-se descrito seguidamente.

Modelo 1:

Este modelo foi construído com recurso às variáveis que se revelaram estatisticamente significativas na construção dos modelos de regressão logística simples anteriormente apresentados. Concretamente, foram descartadas as variáveis que apresentaram um valor-p do teste de Wald superior a 0.15 em todas as categorias possíveis, ou seja, das 22 inicialmente indicadas na Tabela 4.10, foram desconsideradas as variáveis TOTDIAS e BIRTH_WGT.

Modelo 2:

Da análise ao valor-p obtido para cada variável e respetivas categorias do Modelo 1, observa-se que as variáveis SEXO, PRE_OP, MOT_TRANF, DIA_ADM, PROC_3, PROC_4 e PROC_5 não são relevantes estatisticamente quando colocadas num modelo de regressão logística múltiplo conjuntamente com as variáveis selecionadas através da análise aos modelos de regressão logística simples, já que o seu valor-p é superior a 0.05 em todas as suas categorias. Assim, as variáveis referidas foram eliminadas e com as restantes 12 variáveis foi construído o Modelo 2.

Modelo 3:

Da análise ao valor-p obtido para cada variável e respetivas categorias do Modelo 2, observa-se que a variável PROC_2 não é relevante estatisticamente, já que o seu valor-p é superior a 0.05 em todas as suas categorias. Assim, a variável referida foi eliminada e com as restantes 11 variáveis foi construído o Modelo 3.

Modelos 4, 5 e 6:

Utilizando as mesmas 19 variáveis usadas para a construção do Modelo 1, foram construídos 3 modelos com recurso ao método *Stepwise* implementado no *Rstudio* com recurso aos comandos:

- `stepAIC(Modelo1, direction="backward")` para um modelo *backward*,
- `stepAIC(Modelo1, direction="forward")` para um modelo *forward*,
- `stepAIC(Modelo1, direction="both")` para um modelo *both*.

Note-se que é necessária a biblioteca MASS para a utilização dos comandos anteriores.

Na tentativa de encontrar melhores modelos, sugeriu-se considerar as variáveis qualitativas com mais de 2 categorias transformadas em variáveis binárias. Por exemplo, a variável DSP tem 5 categorias possíveis: 1, 2, 6, 7 e 20. A variável é transformada em 5 novas variáveis binárias: DSP_1, DSP_2, DSP_6, DSP_7 e DSP_20, concretamente, por exemplo, se existe uma observação codificada com 6 na variável DSP, esse indivíduo ser-lhe-á atribuído o valor 1 em DSP_6 e 0 nas restantes:

	DSP		DSP 1	DSP 2	DSP 6	DSP 7	DSP 20
Indivíduo 1	2	→	0	1	0	0	0
Indivíduo 2	7		0	0	0	1	0
Indivíduo 3	6		0	0	1	0	0

Tabela 5.19: Exemplo ilustrativo da transformação utilizada

Entenda-se que este processo difere do processo de construção de variáveis *dummy*. Nas variáveis *dummy*, a categoria utilizada como classe de referência é sempre a mesma, isto é, cada categoria de uma dada variável é confrontada com a categoria de referência fixa. Na transformação proposta, cada categoria é agora uma variável binária, pelo que a categoria de referência será sempre alterada. Visitemos o exemplo anterior, focando apenas nas categorias 1, 2 e 6:

Variáveis <i>Dummy</i>		Transformação Proposta		
DSP 1	Classe de referência	DSP 1	0	Classe de referência
DSP 2			1	
DSP 6		DSP 2	0	Classe de referência
			1	
		DSP 6	0	Classe de referência
			1	

Tabela 5.20: Ilustração da diferença entre a utilização de variáveis *dummy* e a transformação proposta

Depois de aplicar a transformação descrita à base de dados, foram construídos os modelos que se descrevem seguidamente.

Modelo 7:

Este modelo, à semelhança do Modelo 1, foi construído com as variáveis que demonstraram significância estatística nos modelos de regressão logística simples anteriormente apresentados, isto é, foram retiradas da base de dados as variáveis TOTDIAS e BIRTH_WGT pois o seu valor-p era superior a 0.15 e procedeu-se à transformação das variáveis tal como descrito. Após a transformação, passou-se de 19 para 204 variáveis, com estas 204 variáveis construiu-se este modelo. Apesar de se ter verificado que neste modelo estão em consideração as mesmas características que as consideradas no Modelo 1, os coeficientes para cada variável são distintos, assim como o valor-p associado. Desta forma, optou-se por apresentar também este modelo.

Modelo 8:

Da análise ao valor-p obtido para cada variável do Modelo 7, foram descartadas as variáveis que obtiveram um valor-p superior a 0.05, ficando as seguintes variáveis: MORF_TUM, ICU, 'DSP 1', 'DSP 6', 'DSP 7', 'DSP 20', 'MOT_TRANF 4', 'DIA_ADM 13', 'DIA_ADM 15', 'DIA_ADM 18', 'DIA_1 5', 'DIA_1 6', 'DIA_1 7', 'DIA_1 8', 'DIA_1 10', 'DIA_1 13', 'DIA_1 14', 'DIA_1 18', 'DIA_2 2', 'DIA_2 15', 'DIA_2 16', 'DIA_2 18', 'DIA_3 2', 'DIA_3 5', 'DIA_3 18', 'DIA_4 5', 'DIA_4 18', 'DIA_5 4', 'DIA_5 18', 'PROC_1 3', 'PROC_1 5', 'PROC_1 6', 'PROC_1 7', 'PROC_1 8', 'PROC_1 9', 'PROC_1 10', 'PROC_1 11', 'PROC_1 12', 'PROC_1 13', 'PROC_1 14', 'PROC_1 15', 'PROC_1 16', 'PROC_1 17', 'PROC_1 18', 'PROC_2 3', 'PROC_2 18', 'PROC_3 18', 'PROC_4 18' e 'PROC_5 18'. Ficou-se assim com 54 variáveis e com elas foi construído este modelo.

Modelo 9:

Da análise ao valor-p obtido para cada variável do Modelo 8, foram descartadas as variáveis que obtiveram um valor-p superior a 0.05. Assim, foram eliminadas as seguintes variáveis: 'DSP 1', 'MOT_TRANF 4', 'DIA_ADM 18', 'DIA_1 18', 'DIA_2 18', 'DIA_3 15', 'DIA_3 18', 'DIA_4 5', 'DIA_4 18', 'DIA_5 18', 'PROC_2 3', 'PROC_2 18', 'PROC_3 18' e 'PROC_4 18'. Diminuiu-se assim o número de variáveis passando de 54 para 40. Com as restantes 40 variáveis foi construído este modelo.

Modelos 10, 11 e 12:

Utilizando as mesmas variáveis usadas para a construção do Modelo 7, foram construídos 3 modelos com recurso ao método *stepwise* implementado no *Rstudio* com recurso aos comandos

- `stepAIC(Modelo7, direction="backward")` para um modelo *backward*,
- `stepAIC(Modelo7, direction="forward")` para um modelo *forward*,
- `stepAIC(Modelo7, direction="both")` para um modelo *both*.

Note-se que é necessária a biblioteca para a utilização dos comandos anteriores.

5.4 COMPARAÇÃO DOS MODELOS

Com o objetivo de escolher o melhor modelo dos obtidos foi realizada uma comparação onde são avaliadas as qualidades dos modelos e as suas capacidades preditivas.

	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6
AIC	10344.05	10294.72	10286.95	10269.43	10344.05	10269.43
BIC	11826.43	11280.32	11163.97	10975.70	11826.43	10975.70
AUC	69%	69.5%	69.5%	69.2%	69%	69.2%
Precisão	68.84%	66.04%	69.82%	74.18%	68.84%	74.18%
Sensibilidade	69.96%	66.34%	71%	76.56%	69.96%	76.56%
Especificidade	58.43%	63.26%	58.84%	52.07%	58.43%	52.07%
Teste de Hosmer & Lemeshow						
χ^2	3.7117	7.9744	7.5752	12.473	3.7117	12.473
df	8	8	8	8	8	8
<i>valor-p</i>	0.8822	0.436	0.476	0.1313	0.8822	0.1313

Tabela 5.21: Medidas de Avaliação dos Modelos 1 a 6

	Modelo 7	Modelo 8	Modelo 9	Modelo 10	Modelo 11	Modelo 12
AIC	10344.05	10267.16	10265.43	10181.67	10344.05	10180.95
BIC	11826.43	10678.51	10568.12	10771.52	11826.43	10778.56
AUC	69%	69.3%	69%	69.7%	69%	69.7%
Precisão	68.84%	59.58%	62%	69.7%	68.84%	74.10%
Sensibilidade	69.96%	58.49%	61.44%	70.81%	69.96%	76.25%
Especificidade	58.43%	69.75%	67.27%	59.39%	58.43%	54.14%
Teste de Hosmer & Lemeshow						
χ^2	3.7117	4.5612	7.2485	11.571	3.7117	6.2683
df	8	8	8	8	8	8
<i>valor-p</i>	0.8822	0.8033	0.1714	0.1714	0.8822	0.6172

Tabela 5.22: Medidas de Avaliação dos Modelos 7 a 12

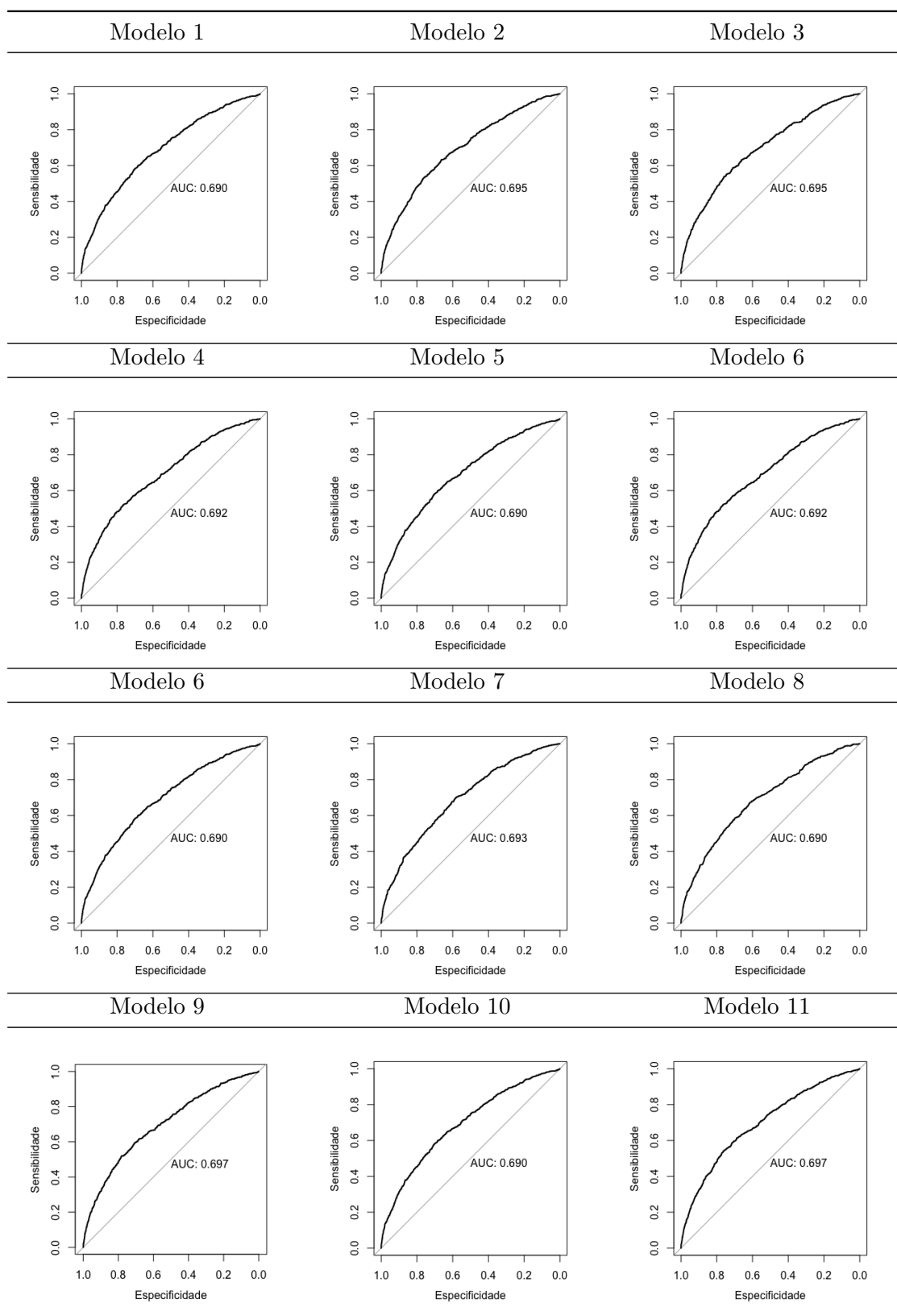


Figura 5.1: Curvas ROC e respectivo AUC dos Modelos 1 a 12

Através de uma análise às tabelas anteriores, conclui-se que nenhum modelo se destaca notoriamente melhor do que os restantes. Segundo o critério AIC o melhor modelo seria o Modelo 12, o mesmo modelo destaca-se também no critério BIC; pelo critério da curva ROC seria o Modelo 10 ou o Modelo 12; na precisão e na sensibilidade destacam-se o Modelo 4 e 6, na especificidade o Modelo 8 tem melhores resultados, enquanto que o valor-p mais elevado do teste Hosmer & Lemeshow pertence aos Modelos 1, 5, 7 e 11.

Salienta-se ainda, que os Modelos 1, 5, 7 e 11 apresentam as mesmas características para os vários critérios de avaliação. Os Modelos 5 e 11 foram obtidos através do método *stepwise* implementado no *Rstudio* com a direção *forward*, sendo que no Modelo 11 foi aplicada a transformação de variáveis explicada anteriormente. Já os Modelos 1 e 7 foram construídos sem nenhuma técnica de *stepwise*, a diferença entre eles é que no Modelo 7 foi aplicada a transformação de variáveis explicada anteriormente.

Uma vez que, o objetivo final passava por desenvolver uma aplicação de fácil utilização e de rápido preenchimento, foi tido em conta o número de variáveis dos modelos.

	Número de Variáveis
Modelo 1	19
Modelo 2	13
Modelo 3	12
Modelo 4	10
Modelo 5	19
Modelo 6	10
Modelo 7	204
Modelo 8	54
Modelo 9	40
Modelo 10	74
Modelo 11	204
Modelo 12	75

Tabela 5.23: Número de variáveis e variáveis de cada modelo

Desta forma, os modelos 7, 8, 9, 10, 11 e 12 não serão considerados para a seleção do melhor modelo, uma vez que embora tenham sido construídos usando um procedimento extensível à construção de variáveis *dummy*, conduziu a um número maior de variáveis binárias e nenhum destes seis modelos apresentou melhores valores em todas as medidas de avaliação consideradas. Analisando de novo a Tabela 5.21, nota-se que os modelos que se destacam mais são o Modelo 3 e o Modelo 4. Optou-se pela escolha do Modelo 3 (com 12 variáveis), que, apesar de ter um valor maior no critério AIC e no critério BIC, apresenta um valor maior na área da curva ROC. Através da análise dos coeficientes associados ao Modelo 3 (Tabela 3, em Anexo), é notório que existem variáveis que afetam significativamente o logaritmo da chance de ser um episódio de reinternamento. A variável MORF_TUM diminui o risco de ser reinternamento, isto é, caso o doente tenha um tumor, o parâmetro que entrará para o cálculo será -1.171, já na variável DSP tem-se que a categoria 7 aumenta o risco pois o

seu parâmetro é 1.520. Na variável DIA__1 a categoria 15 é a que afeta mais o risco de ser reinternamento tendo como parâmetro 2.063. A categoria 15 das variáveis DIA__2 e DIA__3 diminui o risco de se tratar de um episódio de reinternamento com um parâmetro igual a -0.908 e -1.482 respectivamente. A variável PROC__1 diminui o risco de ser reinternamento independentemente da categoria escolhida pois os seus parâmetros são todos negativos.

Capítulo 6

Aplicação Shiny

A forma como é apresentado o modelo e a previsão do risco de ser um episódio de reinternamento obtida é muito importante. Deve ser claro, conciso e simples.

Com esse objetivo, recorreu-se ao *Shiny* - sistema para o desenvolvimento de aplicações *web* - conjuntamente com o *software Rstudio*. Este conjunto permite que qualquer utilizador consiga interpretar o resultado e tomar decisões mais conscientes.

Desenvolveu-se uma aplicação *web* que calcula a previsão de um dado episódio ser de reinternamento para cada doente em qualquer momento e de forma imediata.

Após o preenchimento das variáveis presentes no Modelo 3, é calculada a previsão do risco de se tratar de um reinternamento, de acordo com esse modelo. A cor apresentada é representativa do risco associado, sendo que a cor vermelha representa risco elevado de ser reinternamento (valores entre 75% e 100%), a cor amarela representa risco moderado (valores entre 50% e 74%) e a cor verde representa risco baixo de ser reinternamento (valores entre 0% e 49%). De seguida, é possível visualizar um caso para cada um dos riscos.

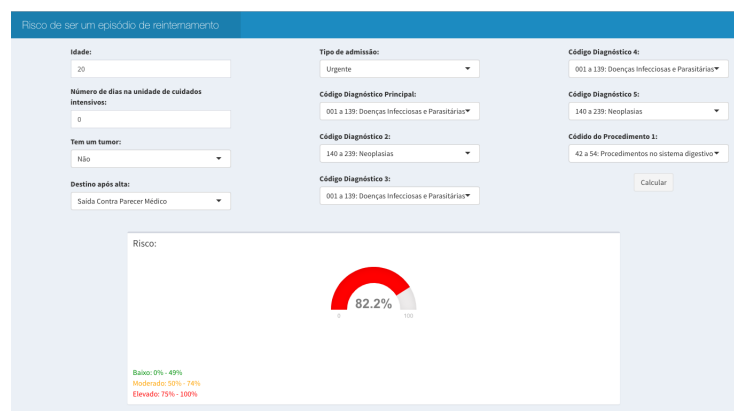


Figura 6.1: Exemplo de um caso com risco elevado de se tratar de um episódio de reinternamento

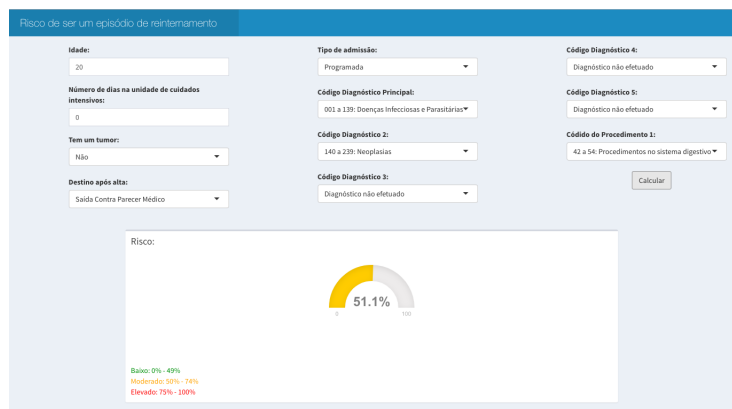


Figura 6.2: Exemplo de um caso com risco moderado de se tratar de um episódio de reinternamento

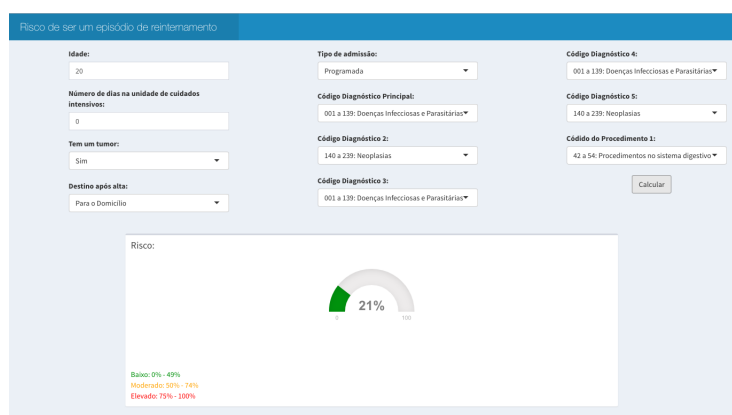


Figura 6.3: Exemplo de um caso com risco baixo de se tratar de um episódio de reinternamento

De referir que os campos de preenchimento, "Idade" e "Número de dias na unidade de cuidados intensivos" apenas admitem caracteres numéricos. Ao introduzir outro tipo de carácter a célula aparece vermelha. Nos restantes campos apenas é permitido escolher uma opção.

Figura 6.4: Exemplo de mau preenchimento na aplicação *Shiny*

Existem grupos de códigos de procedimentos que não se encontram disponíveis para escolha, uma vez que o modelo usado não contém coeficientes para os mesmos, nomeadamente para os grupos: 0 - "Procedimentos e intervenções não classificadas noutra parte", 1 - "Procedimentos no sistema nervoso", 2 - "Procedimentos no sistema endócrino" e 4 - "Outros procedimentos de diagnóstico e terapêutica não classificados noutra parte".

Conclusões

Ao longo deste estudo várias dificuldades foram sentidas. Inicialmente começou-se por estudar Modelação por Caminhos (*Path Modeling*), e foi o objeto de estudo por mais de um mês. Foi necessário um igual período de tempo para o estudo dos conceitos do financiamento hospitalar. Quando se recebeu a base de dados com a qual se iria trabalhar, tal como foi descrito anteriormente, surgiram alguns problemas. Percebeu-se que a base de dados não permitia o uso da metodologia estudada, pois a maioria das variáveis presentes eram qualitativas. Foi sugerido, por parte da empresa, que se calculasse a matriz de correlação das várias variáveis. Após o estudo sobre quais seriam as melhores medidas de associação para o tipo de variáveis presentes, fez-se um *script* no *Rstudio* que calculasse a matriz. A par disso percebeu-se que os hospitais com taxas de reinternamento elevadas são penalizados e foi esse o caminho que se seguiu: modelar o reinternamento podendo prever se um doente seria reinternado.

De início, codificou-se a variável de reinternamento, mas não se teve em consideração o tempo entre episódios, o que fez com que o modelo obtido não estivesse ajustado aos dados. Então voltou-se a considerar a variável codificada pela empresa, reduzindo drasticamente o número de observações de 184 267 para 24 782. Ao considerar esta variável, deixou de ser possível construir um modelo que previsse se um doente seria ou não reinternado. Assim, modelou-se a variável que previa se um dado episódio seria ou não um reinternamento. Este estudo permite entender quais as variáveis que influenciam mais o risco de ser ou não um episódio de reinternamento, tornando possível o esclarecimento de algumas relações que se podem revelar interessantes.

Com a matriz de medidas de associação calculada, concluiu-se que não existiam variáveis fortemente relacionadas com RINT, podendo ser possível considerar todas as variáveis para o modelo de regressão. Como a variável resposta, RINT, é binária, recorreu-se ao modelo de regressão logística. Inicialmente, foram construídos e analisados os modelos de regressão logística simples, um modelo para cada variável explicativa. As variáveis preditoras que não revelaram interesse para o estudo foram deixadas de parte e passou-se para um modelo de regressão logística múltipla. Foram feitos 12 modelos distintos:

- 6 modelos foram feitos com as variáveis originais, sendo que 3 foram obtidos por três vias distintas: *forward*, *backward* e *both* do método *Stepwise*,
- após a transformação das variáveis categóricas com mais de 2 categorias em variáveis binárias foram feitos 6 modelos, sendo que 3 foram obtidos por três vias distintas: *forward*, *backward* e *both* do método *Stepwise*,

A transformação das variáveis categóricas com mais de 2 categorias em variáveis binárias foi benéfica no sentido que se obteve um modelo com melhores capacidades preditivas, o Modelo

12. Contudo, esta transformação é computacionalmente mais pesada, isto é, inicialmente trabalhou-se com 19 variáveis preditoras e após esta transformação passou-se a trabalhar com 204. O tempo de cálculo do *Rstudio* para a obtenção dos modelos 10, 11 e 12 foi muito maior comparativamente ao tempo com que se obtiveram os modelos 4, 5 e 6. Por exemplo, o modelo 10 demorou cerca de 18h enquanto que o modelo 4 foi calculado em pouco mais de 1h. Para cada modelo foi calculado o AIC, o BIC, a precisão, a sensibilidade, a especificidade, fez-se a curva ROC e foi realizado o teste de Hosmer & Lemeshow para avaliar o ajustamento do modelo aos dados. Depois de analisar e comparar as medidas de avaliação de todos os modelos, concluiu-se que nenhum modelo se destacava em todos os critérios. Teve-se também em consideração o número de variáveis de cada modelo. Por essa razão os modelos 7, 8, 9, 10, 11 e 12 foram excluídos da seleção do melhor modelo, pois apresentavam um elevado número de variáveis o que tornava a utilização da aplicação desenvolvida demorada e mais complexa. Os modelos 4 e 6 destacaram-se nos critérios AIC e BIC, na precisão e na sensibilidade. O Modelo 3 destacou-se na curva ROC e o Modelo 2 na especificidade. Os modelos 1 e 5 têm o maior valor-p do teste de Hosmer e Lemeshow. Os modelos 4 e 6 são os que se destacam em maior número de medidas de classificação, porém são os modelos com menor número de variáveis, podendo o modelo estar sobreajustado aos dados, isto é, o modelo pode estar a ficar demasiado ajustado aos dados e perder capacidades preditivas. Assim, foi selecionado o Modelo 3.

Tabela 7.1: Significância estatística das variáveis presentes no modelo selecionado: IC, intervalo de confiança; OR, Razão de chances (*Odds Ratio*); AOR, Razão de chances ajustada (*Adjusted Odds Ratio*).

	Análise Simples		Análise Múltipla
	OR (IC a 95%)	valor-p	AOR (IC a 95%)
Morfologia Tumoral:	Classe de referência		
Sem tumor			
Com tumor	0.412 [0.297; 0.572]	<0.001	0.310 [0.201; 0.479]
Destino após alta:	Classe de referência		
Domicílio			
Outra instituição com internamento	1.458 [1.108; 1.917]	0.007	1.391 [0.995; 1.944]
Serviço domiciliário	0.758 [0.528; 1.089]	0.134	0.605 [0.378; 0.970]
Saída contra parecer médico	5.437 [4.273; 6.916]	<0.001	4.572 [3.342; 6.256]
Falecimento	3.209 [2.768; 3.721]	<0.001	2.645 [2.167; 3.227]
ICU	0.969 [0.945; 0.994]	0.016	0.970 [0.942; 0.998]
Tipo de admissão:	Classe de referência		
Programada			
Urgente	1.358 [1.239; 1.488]	<0.001	1.177 [1.027; 1.348]
SIGIC	0.219 [0.126; 0.381]	<0.001	0.420 [0.220; 0.800]

Continuação da Tabela 7.1

	Análise Simples		Análise Múltipla
	OR (IC a 95%)	valor-p	AOR (IC a 95%)
AGE	0.994 [0.993; 0.996]	<0.001	0.993 [0.990; 0.995]
Diagnóstico Principal:			
Doenças Infecciosas e Parasitárias	0.518 [0.186; 1.443]	0.208	0.841 [0.282; 2.509]
Neoplasias	Classe de referência		
Doenças Endócrinas, Nutricionais, Metabólicas e da Imunidade	0.583 [0.451; 0.754]	<0.001	0.766 [0.558; 1.051]
Doenças do Sangue e dos Órgãos Hematopoiéticos	0.838 [0.515; 1.364]	0.477	0.675 [0.363; 1.256]
Doenças Mentais	0.363 [0.232; 0.570]	<0.001	0.346 [0.201; 0.597]
Doenças do Sistema Nervoso e dos Órgãos dos Sentidos	0.384 [0.301; 0.489]	<0.001	0.409 [0.297; 0.564]
Doenças do Aparelho Circulatório	0.458 [0.394; 0.532]	<0.001	0.496 [0.401; 0.613]
Doenças do Aparelho Respiratório	0.631 [0.541; 0.736]	<0.001	0.597 [0.484; 0.736]
Doenças do Aparelho Digestivo	0.766 [0.656; 0.895]	<0.001	0.685 [0.555; 0.845]
Doenças do Aparelho Geniturinário	0.449 [0.367; 0.551]	<0.001	0.518 [0.397; 0.677]
Complicações da Gravidez, do Parto e do Puerpério	0.843 [0.659; 1.078]	0.173	0.911 [0.563; 1.473]
Doenças da Pele e do Tecido Subcutâneo	0.494 [0.289; 0.842]	0.010	0.571 [0.289; 1.127]
Doenças do Aparelho Osteomuscular e do Tecido Conjuntivo	0.383 [0.271; 0.540]	<0.001	0.738 [0.485; 1.123]
Anomalias Congénitas	0.349 [0.239; 0.510]	<0.001	0.466 [0.295; 0.735]
Certas Condições originadas no Período Perinatal	7.226 [4.350; 12.005]	<0.001	7.871 [3.744; 16.547]
Sintomas, Sinais e Condições Mal Definidas	0.698 [0.473; 1.029]	0.069	0.478 [0.283; 0.808]
Lesões e Intoxicações	1.005 [0.863; 1.169]	0.953	1.199 [0.964; 1.492]
Classificação Suplementar Factores com Influência no Estado de Saúde e Contacto com Serviços Saúde	0.762 [0.623; 0.932]	0.008	0.913 [0.687; 1.213]
Diagnóstico Secundário 2:			
Diagnóstico não efetuado	Classe de referência		
Doenças Infecciosas e Parasitárias	1.028 [0.509; 2.075]	0.939	0.785 [0.318; 1.939]
Neoplasias	2.358 [1.803; 3.083]	<0.001	2.025 [1.370; 2.995]
Doenças Endócrinas, Nutricionais, Metabólicas e da Imunidade	0.787 [0.620; 0.998]	0.048	1.032 [0.719; 1.481]
Doenças do Sangue e dos Órgãos Hematopoiéticos	1.937 [1.491; 2.515]	<0.001	1.563 [1.057; 2.312]

Continuação da Tabela 7.1

	Análise Simples		Análise Múltipla
	OR (IC a 95%)	valor-p	AOR (IC a 95%)
Doenças Mentais	0.971 [0.736; 1.282]	0.838	1.141 [0.764; 1.705]
Doenças do Sistema Nervoso e dos Órgãos dos Sentidos	1.180 [0.902; 1.544]	0.227	1.210 [0.816; 1.794]
Doenças do Aparelho Circulatório	0.871 [0.709; 1.071]	0.189	1.254 [0.903; 1.742]
Doenças do Aparelho Respiratório	1.009 [0.757; 1.346]	0.951	1.189 [0.795; 1.778]
Doenças do Aparelho Digestivo	1.226 [0.958; 1.569]	0.106	1.116 [0.771; 1.617]
Doenças do Aparelho Geniturinário	0.906 [0.701; 1.170]	0.450	1.083 [0.743; 1.578]
Complicações da Gravidez, do Parto e do Puerpério	1.228 [0.735; 2.052]	0.433	0.800 [0.354; 1.804]
Doenças da Pele e do Tecido Subcutâneo	1.202 [0.818; 1.765]	0.349	1.215 [0.737; 2.004]
Doenças do Aparelho Osteomuscular e do Tecido Conjuntivo	0.876 [0.611; 1.255]	0.469	1.178 [0.731; 1.897]
Anomalias Congénitas	0.850 [0.549; 1.317]	0.467	1.200 [0.700; 2.057]
Certas Condições originadas no Período Perinatal	0.436 [0.220; 0.865]	0.018	0.403 [0.172; 0.945]
Sintomas, Sinais e Condições Mal Definidas	1.782 [1.459; 2.177]	<0.001	1.539 [1.113; 2.128]
Lesões e Intoxicações	1.237 [0.901; 1.699]	0.188	1.222 [0.784; 1.905]
Classificação Suplementar Factores com Influência no Estado de Saúde e Contacto com Serviços Saúde	1.070 [0.895; 1.279]	0.458	1.232 [0.914; 1.661]
Diagnóstico Secundário 3:	Classe de referência		
Diagnóstico não efetuado			
Doenças Infecciosas e Parasitárias	2.096 [1.266; 3.470]	0.004	1.590 [0.857; 2.950]
Neoplasias	2.074 [1.586; 2.712]	<0.001	1.528 [1.026; 2.274]
Doenças Endócrinas, Nutricionais, Metabólicas e da Imunidade	0.943 [0.769; 1.156]	0.572	0.999 [0.715; 1.396]
Doenças do Sangue e dos Órgãos Hematopoiéticos	1.764 [1.396; 2.228]	<0.001	1.199 [0.826; 1.742]
Doenças Mentais	0.730 [0.547; 0.975]	0.033	0.757 [0.500; 1.146]
Doenças do Sistema Nervoso e dos Órgãos dos Sentidos	1.005 [0.781; 1.292]	0.971	1.218 [0.843; 1.760]
Doenças do Aparelho Circulatório	0.858 [0.727; 1.013]	0.071	0.927 [0.683; 1.258]
Doenças do Aparelho Respiratório	1.034 [0.809; 1.320]	0.790	1.007 [0.697; 1.454]
Doenças do Aparelho Digestivo	1.438 [1.167; 1.772]	<0.001	1.228 [0.876; 1.721]
Doenças do Aparelho Geniturinário	1.079 [0.862; 1.351]	0.507	0.949 [0.659; 1.366]

Continuação da Tabela 7.1

	Análise Simples		Análise Múltipla
	OR (IC a 95%)	valor-p	AOR (IC a 95%)
Complicações da Gravidez, do Parto e do Puerpério	1.303 [0.826; 2.055]	0.255	0.642 [0.290; 1.424]
Doenças da Pele e do Tecido Subcutâneo	1.523 [1.112; 2.088]	0.009	1.423 [0.916; 2.211]
Doenças do Aparelho Osteomuscular e do Tecido Conjuntivo	0.880 [0.610; 1.269]	0.493	1.033 [0.641; 1.666]
Anomalias Congénitas	0.743 [0.455; 1.212]	0.234	0.730 [0.388; 1.375]
Certas Condições originadas no Período Perinatal	0.456 [0.185; 1.123]	0.088	0.227 [0.057; 0.897]
Sintomas, Sinais e Condições Mal Definidas	1.554 [1.318; 1.832]	<0.001	1.086 [0.802; 1.470]
Lesões e Intoxicações	1.384 [1.048; 1.827]	0.022	1.022 [0.671; 1.556]
Classificação Suplementar Factores com Influência no Estado de Saúde e Contacto com Serviços Saúde	1.078 [0.934; 1.244]	0.305	1.063 [0.805; 1.405]
Diagnóstico Secundário 4:	Classe de referência		
Diagnóstico não efetuado			
Doenças Infecciosas e Parasitárias	1.617 [0.897; 2.916]	0.110	1.731 [0.846; 3.542]
Neoplasias	2.059 [1.563; 2.713]	<0.001	1.268 [0.835; 1.925]
Doenças Endócrinas, Nutricionais, Metabólicas e da Imunidade	0.943 [0.775; 1.148]	0.559	0.980 [0.699; 1.374]
Doenças do Sangue e dos Órgãos Hematopoiéticos	1.920 [1.526; 2.417]	<0.001	1.485 [1.027; 2.148]
Doenças Mentais	0.898 [0.675; 1.196]	0.462	0.838 [0.555; 1.267]
Doenças do Sistema Nervoso e dos Órgãos dos Sentidos	1.034 [0.797; 1.342]	0.800	0.924 [0.609; 1.401]
Doenças do Aparelho Circulatório	0.933 [0.803; 1.084]	0.363	1.045 [0.774; 1.412]
Doenças do Aparelho Respiratório	1.203 [0.952; 1.522]	0.122	0.933 [0.630; 1.382]
Doenças do Aparelho Digestivo	1.366 [1.101; 1.695]	0.005	1.140 [0.795; 1.635]
Doenças do Aparelho Geniturinário	1.052 [0.833; 1.329]	0.669	0.950 [0.647; 1.393]
Complicações da Gravidez, do Parto e do Puerpério	1.149 [0.644; 2.050]	0.638	0.895 [0.346; 2.317]
Doenças da Pele e do Tecido Subcutâneo	1.301 [0.911; 1.859]	0.148	0.935 [0.567; 1.542]
Doenças do Aparelho Osteomuscular e do Tecido Conjuntivo	0.923 [0.617; 1.380]	0.695	0.878 [0.498; 1.548]
Anomalias Congénitas	0.698 [0.387; 1.259]	0.232	0.750 [0.342; 1.644]
Certas Condições originadas no Período Perinatal	0.733 [0.295; 1.820]	0.503	1.135 [0.308; 4.178]

Continuação da Tabela 7.1

	Análise Simples		Análise Múltipla
	OR (IC a 95%)	valor-p	AOR (IC a 95%)
Sintomas, Sinais e Condições Mal Definidas	1.602 [1.372; 1.871]	<0.001	1.285 [0.948; 1.743]
Lesões e Intoxicações	1.544 [1.146; 2.081]	0.004	1.404 [0.897; 2.196]
Classificação Suplementar Factores com Influência no Estado de Saúde e Contacto com Serviços Saúde	1.149 [1.002; 1.317]	0.046	1.230 [0.924; 1.636]
Diagnóstico Secundário 5:	Classe de referência		
Diagnóstico não efetuado	1.387 [0.714; 2.696]	0.3343	1.016 [0.436; 2.367]
Doenças Infecciosas e Parasitárias	1.846 [1.379; 2.471]	<0.001	1.364 [0.915; 2.034]
Neoplasias	0.892 [0.735; 1.082]	0.246	0.917 [0.679; 1.239]
Doenças Endócrinas, Nutricionais, Metabólicas e da Imunidade	2.111 [1.706; 2.613]	<0.001	1.645 [1.188; 2.279]
Doenças do Sangue e dos Órgãos Hematopoiéticos	1.218 [0.937; 1.584]	0.140	1.316 [0.917; 1.889]
Doenças Mentais	0.952 [0.730; 1.241]	0.715	0.906 [0.619; 1.325]
Doenças do Sistema Nervoso e dos Órgãos dos Sentidos	0.874 [0.755; 1.012]	0.072	1.103 [0.856; 1.421]
Doenças do Aparelho Circulatório	1.188 [0.940; 1.501]	0.150	1.153 [0.823; 1.614]
Doenças do Aparelho Respiratório	1.417 [1.141; 1.759]	0.002	1.088 [0.785; 1.507]
Doenças do Aparelho Digestivo	1.153 [0.916; 1.452]	0.224	1.212 [0.871; 1.687]
Doenças do Aparelho Geniturinário	1.467 [0.797; 2.698]	0.218	1.189 [0.460; 3.072]
Complicações da Gravidez, do Parto e do Puerpério	1.467 [1.021; 2.106]	0.038	1.419 [0.901; 2.234]
Doenças da Pele e do Tecido Subcutâneo	0.753 [0.485; 1.168]	0.205	0.869 [0.499; 1.516]
Doenças do Aparelho Osteomuscular e do Tecido Conjuntivo	0.697 [0.377; 1.290]	0.250	0.682 [0.312; 1.493]
Anomalias Congénitas	0.411 [0.100; 1.690]	0.218	0.407 [0.078; 2.125]
Certas Condições originadas no Período Perinatal	1.539 [1.314; 1.802]	<0.001	1.082 [0.827; 1.415]
Sintomas, Sinais e Condições Mal Definidas	1.723 [1.261; 2.353]	<0.001	1.382 [0.891; 2.144]
Lesões e Intoxicações	0.943 [0.814; 1.093]	0.434	0.966 [0.752; 1.242]
Classificação Suplementar Factores com Influência no Estado de Saúde e Contacto com Serviços Saúde			
Procedimento:			
Procedimentos no olho	0.120 [0.070; 0.205]	<0.001	0.149 [0.078; 0.285]

Continuação da Tabela 7.1

	Análise Simples		Análise Múltipla
	OR (IC a 95%)	valor-p	AOR (IC a 95%)
Outros procedimentos de diagnóstico e terapêutica não classificados noutra parte	0.440 [0.245; 0.793]	< 0.001	0.293 [0.142; 0.606]
Procedimentos no nariz, boca e faringe	0.239 [0.136; 0.420]	< 0.001	0.189 [0.092; 0.389]
Procedimentos no sistema respiratório	0.339 [0.216; 0.533]	< 0.001	0.257 [0.147; 0.447]
Procedimentos no sistema cardiovascular	0.255 [0.170; 0.384]	< 0.001	0.284 [0.174; 0.464]
Procedimentos no sistema linfático e imunológico	0.430 [0.243; 0.762]	< 0.001	0.292 [0.147; 0.578]
Procedimentos no sistema digestivo	0.462 [0.311; 0.685]	< 0.001	0.388 [0.242; 0.621]
Procedimentos do sistema urinário	0.241 [0.155; 0.375]	< 0.001	0.287 [0.169; 0.487]
Procedimentos nos órgãos genitais masculinas	0.142 [0.074; 0.273]	< 0.001	0.185 [0.082; 0.416]
Procedimentos nos órgãos genitais femininas	0.370 [0.219; 0.626]	< 0.001	0.309 [0.163; 0.586]
Procedimentos obstétricos	0.419 [0.259; 0.680]	< 0.001	0.260 [0.125; 0.542]
Procedimentos no sistema musculoesquelético	0.197 [0.127; 0.307]	< 0.001	0.158 [0.092; 0.271]
Procedimentos no sistema integumentário	0.234 [0.144; 0.380]	< 0.001	0.167 [0.093; 0.299]
Outros procedimentos de diagnóstico e terapêutica não classificados	0.461 [0.316; 0.673]	< 0.001	0.354 [0.226; 0.556]
Procedimento não efetuado	Classe de referência		

De acordo com o AOR significativo, do modelo estimado, conclui-se que o risco de se tratar de um episódio de reinternamento é menor nos doentes:

- com tumor, do que nos doentes sem tumor (AOR=0.310).
- que seguiram para um serviço domiciliário após a alta, comparativamente aos doentes que foram para o seu domicílio (AOR=0.605).
- com maior número de dias na unidade dos cuidados intensivos (AOR=0.970).
- com maior idade (AOR=0.993).
- admitidos no tipo SIGIC comparativamente aos doentes com admissão programada (AOR=0.420).
- com diagnóstico principal de doença mental (AOR=0.346), doenças do sistema nervoso e dos órgãos dos sentidos (AOR=0.409), do aparelho circulatório (AOR=0.496), do aparelho respiratório (AOR=0.597), do aparelho digestivo (AOR=0.685), do aparelho geniturinário (AOR=0.518), com anomalias congénitas (AOR=0.466) ou com sintomas, sinais e condições mal definidas (AOR=0.478) comparativamente a doentes cujo diagnóstico principal é neoplasias.
- com diagnóstico secundário 2 de certas condições originadas no período perinatal (AOR=0.403) comparativamente a doentes sem segundo diagnóstico.

- com diagnóstico secundário 3 de certas condições originadas no período perinatal (AOR=0.227) comparativamente a doentes sem terceiro diagnóstico.
- que são sujeitos a algum tipo de procedimento comparativamente aos doentes que não são submetidos a nenhum procedimento.

Mas é maior o risco de ser um reinternamento nos doentes:

- com saída contra parecer médico (AOR=4.572) ou por falecimento (AOR=2.645) comparativamente a doentes que seguiram para o seu domicílio após a alta médica.
- que foram admitidos numa urgência (AOR=1.177) comparativamente aos doentes com admissão programada.
- com diagnóstico principal certas condições originadas no período perinatal (AOR=7.871) comparando ao grupo de doentes com diagnóstico principal de neoplasias.
- com diagnóstico secundário de neoplasias (AOR=2.025), doenças de sangue e dos órgãos hematopoiéticos (AOR=1.563) ou com sintomas, sinais e condições mal definidas (AOR=1.539) comparativamente ao grupo de doentes sem segundo diagnóstico.
- com terceiro diagnóstico de neoplasias (AOR=1.538) comparativamente aos doentes sem terceiro diagnóstico efetuado.
- com quarto diagnóstico de doenças no sangue e dos órgãos hematopoiéticos (AOR=1.485) comparativamente ao grupo de doentes sem quarto diagnóstico efetuado.
- com quinto diagnóstico de doenças no sangue (AOR=1.645) comparativamente aos doentes sem quinto diagnóstico efetuado.

Posteriormente, com este modelo foi desenvolvida uma aplicação com recurso ao *Shiny* - sistema para o desenvolvimento de aplicações web - conjuntamente com o software *RStudio*. Esta aplicação permite calcular em tempo real o risco de ser um episódio de reinternamento. O valor é devolvido ao utilizador num gráfico cuja cor é representativa do risco associado, assumindo as cores: vermelho (risco superior a 75%), amarelo e verde (inferior a 49%).

Como trabalho futuro, seria importante fazer uma reapreciação do Modelo 3, proposto com a selecção das categorias de referência definidas com o auxílio de um agente clínico, conhecedor de realidades em ambiente hospitalar, ou ainda redefinir as categorias a considerar, agregando categorias, diminuindo assim, o número de parâmetros a estimar no modelo final (reduzindo a sua complexidade). Por exemplo, na variável PROC_1 todas as categorias obtiveram um AOR significativo, o que indica que o que importa é haver procedimento, e não necessariamente a identificação do tipo de procedimento realizado. No mesmo âmbito, seria interessante a criação de uma nova variável que identificasse os episódios que iriam resultar em reinternamento. Com esse estudo, poderia ser possível comparar as conclusões deste trabalho e analisar se existe alguma semelhança entre o risco de ser reinternamento e o risco de reinternamento. Ainda como proposta de trabalho futuro, poderia ainda ser interessante usar outros métodos para prever se um doente seria ou não reinternado, recorrer, por exemplo, a uma árvore de decisão. Também com o intuito de oferecer maior controlo na gestão hospitalar, poderia ser útil, recorrendo a séries temporais estudar os recursos necessários para cada doente de um dado GDH, evitando a falta e/ou excesso de recursos num determinado mês.

Referências

- [1] R. A. L. da Silva, «Financiamento Hospitalar - Análise e Reflexão», tese de mestrado, Instituto de Ciências Biomédicas Abel Salazar - Universidade do Porto, 2015.
- [2] P. S. O. F. Nunes, «Financiamento Hospitalar em Portugal - Incentivos à Selecção e Equidade», tese de mestrado, Escola Nacional de Saúde Pública - Universidade Nova de Lisboa, 2012.
- [3] C. M. de Sé Reis, «Modelos de Gestão Hospitalar - Análise Comparativa», tese de mestrado, Faculdade de Economia da Universidade de Coimbra, 2011.
- [4] A. Hall, URL: <http://sweet.ua.pt/andreia.hall/TEA/Capcorrel.pdf>.
- [5] D. Cramer, *Fundamental Statistics for Social Research: Step-by-Step Calculations and Computer Techniques Using SPSS for Windows*. Taylor & Francis, 2003.
- [6] S. Vieira, *Introdução à Bioestatística*. Rio de Janeiro : Campus, 1980.
- [7] E. Gonçalves e N. M. Lopes, *Estatística: Teoria Matemática e Aplicações*. Escolar Editora, 2005.
- [8] J. A. Nelder e R. W. M. Wedderburn, «Generalized Linear Models», *Journal of the Royal Statistical Society. Series A (General)*, vol. 135, n.º 3, pp. 370–384, 1972. URL: <http://www.jstor.org/stable/2344614>.
- [9] M. A. A. Turkman e G. L. Silva, *Modelos Lineares Generalizados - da teoria à prática*. Edições SPE, 2000.
- [10] J. Cadima, «Notas de disciplina: Modelos Matemáticos e Aplicações Modelos Lineares Generalizados», 2015-16. URL: [https://fenix.isa.ulisboa.pt/qubEdu/disciplinas/mmapl-4/2015-2016/2-semester/lateral/materiais-de-apoio/modulo-ii-\(modelacao-estatistica\)/textos/modelos-lineares-generalizados](https://fenix.isa.ulisboa.pt/qubEdu/disciplinas/mmapl-4/2015-2016/2-semester/lateral/materiais-de-apoio/modulo-ii-(modelacao-estatistica)/textos/modelos-lineares-generalizados).
- [11] D. W. Hosmer e S. Lemeshow, *Applied Logistic Regression*, 2nd Edition. John Wiley & Sons, 2000.
- [12] J. Marôco, *Análise Estatística com o SPSS.: 3ª edição*. Edições Sílabo, 2007.
- [13] M. H. Pestana e J. N. Gageiro, *Análise de Dados para Ciências Sociais.: 6ª edição*. Edições Sílabo, 2014.
- [14] D. Cox, *Analysis of Binary Data*. CRC Press, 2018.
- [15] P. C. Emiliano, M. J. Vivanco e F. S. de Menezes, «Information criteria: How do they behave in different models?», *Computational Statistics & Data Analysis*, vol. 69, pp. 141–153, 2014. URL: <http://www.sciencedirect.com/science/article/pii/S0167947313002776>.
- [16] P. Tattar, *Statistical Application Development with R and Python*. Packt Publishing, 2017.
- [17] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2017. URL: <https://www.R-project.org/>.

Apêndice A
Tabelas relativas aos modelos construídos

Tabela 1: Coeficientes β , Teste de Wald (p-valor), OR e respectivo IC a 95% relativos ao Modelo 1

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
(Intercept)	-0.858670	0.001362	0.424 [0.251; 0.717]
SEXO2	0.011476	0.839188	1.012 [0.905; 1.130]
MORF_TUM1	-1.097979	1.08e-06	0.334 [0.215; 0.519]
DSP2	0.706691	0.518680	2.027 [0.237; 17.337]
DSP6	-0.475900	0.048785	0.621 [0.387; 0.997]
DSP7	1.499676	<2e-16	4.480 [3.262; 6.153]
DSP20	0.952428	<2e-16	2.592 [2.121; 3.168]
PRE_OP	-0.005044	0.514795	0.995 [0.980; 1.010]
ICU	-0.030007	0.047713	0.970 [0.942; 1.000]
ADM_TIP2	0.184220	0.015884	1.202 [1.035; 1.396]
ADM_TIP6	-0.917847	0.005435	0.399 [0.209; 0.763]
AGE	-0.008070	1.37e-08	0.992 [0.989; 0.995]
MOT_TRANF2	-0.369006	0.733342	0.691 [0.083; 5.777]
MOT_TRANF3	-14.360962	0.980314	0.000 [0.000; Inf]
MOT_TRANF4	0.445369	0.753870	1.561 [0.096; 25.265]
DIA_ADM1	-0.941827	0.386100	0.390 [0.046; 3.281]
DIA_ADM3	-0.309330	0.212147	0.734 [0.451; 1.193]
DIA_ADM4	0.014367	0.967092	1.014 [0.513; 2.008]
DIA_ADM5	-0.134344	0.775682	0.874 [0.347; 2.203]
DIA_ADM6	-0.261509	0.286572	0.770 [0.476; 1.245]
DIA_ADM7	-0.099145	0.596633	0.906 [0.627; 1.307]
DIA_ADM8	-0.107610	0.554547	0.898 [0.628; 1.283]
DIA_ADM9	-0.240909	0.175057	0.786 [0.555; 1.113]
DIA_ADM10	-0.259417	0.157948	0.772 [0.538; 1.106]
DIA_ADM11	0.076825	0.852395	1.080 [0.481; 2.426]
DIA_ADM12	-0.449616	0.311311	0.638 [0.267; 1.523]
DIA_ADM13	0.629972	0.086487	1.878 [0.914; 3.858]
DIA_ADM14	-0.048282	0.900098	0.953 [0.448; 2.025]
DIA_ADM15	1.037104	0.110066	2.821 [0.791; 10.066]
DIA_ADM16	-0.284643	0.065943	0.752 [0.555; 1.019]
DIA_ADM17	-0.098008	0.624833	0.907 [0.612; 1.343]
DIA_ADM18	-0.338046	0.075060	0.713 [0.492; 1.035]
DIA_11	0.106948	0.858691	1.113 [0.343; 3.612]
DIA_13	-0.059743	0.798039	0.942 [0.596; 1.489]
DIA_14	-0.358893	0.366303	0.698 [0.321; 1.522]
DIA_15	-0.959825	0.044045	0.383 [0.150; 0.975]

Continuação da Tabela 1

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
DIA_16	-0.680474	0.002306	0.506 [0.327; 0.784]
DIA_17	-0.651840	0.000258	0.521 [0.367; 0.739]
DIA_18	-0.457089	0.009544	0.633 [0.448; 0.895]
DIA_19	-0.208822	0.219704	0.812 [0.581; 1.133]
DIA_110	-0.461172	0.017035	0.631 [0.432; 0.921]
DIA_111	-0.096496	0.829593	0.908 [0.377; 2.186]
DIA_112	-0.269147	0.561727	0.764 [0.308; 1.896]
DIA_113	-0.841239	0.024921	0.431 [0.207; 0.899]
DIA_114	-0.765213	0.032685	0.465 [0.231; 0.939]
DIA_115	1.180752	0.071712	3.257 [0.901; 11.772]
DIA_116	-0.533398	0.065722	0.587 [0.332; 1.035]
DIA_117	0.270483	0.124631	1.311 [0.928; 1.851]
DIA_118	0.097071	0.616816	1.102 [0.753; 1.612]
DIA_21	-0.235360	0.611759	0.790 [0.318; 1.961]
DIA_22	0.736047	0.000303	2.088 [1.400; 3.112]
DIA_23	0.083616	0.653973	1.087 [0.754; 1.567]
DIA_24	0.476431	0.018724	1.610 [1.082; 2.396]
DIA_25	0.154974	0.454515	1.168 [0.778; 1.753]
DIA_26	0.213949	0.294709	1.239 [0.830; 1.848]
DIA_27	0.260622	0.125980	1.298 [0.929; 1.812]
DIA_28	0.216799	0.298170	1.242 [0.826; 1.869]
DIA_29	0.167398	0.382240	1.182 [0.812; 1.721]
DIA_210	0.141070	0.469470	1.152 [0.786; 1.688]
DIA_211	-0.128872	0.759924	0.879 [0.385; 2.009]
DIA_212	0.237331	0.357644	1.268 [0.765; 2.102]
DIA_213	0.164703	0.502462	1.179 [0.729; 1.908]
DIA_214	0.233629	0.399930	1.263 [0.733; 2.176]
DIA_215	-0.887833	0.044757	0.412 [0.173; 0.979]
DIA_216	0.472678	0.004911	1.604 [1.154; 2.230]
DIA_217	0.210653	0.361199	1.234 [0.785; 1.940]
DIA_218	0.255943	0.098638	1.292 [0.953; 1.750]
DIA_31	0.440229	0.164160	1.553 [0.835; 2.888]
DIA_32	0.413397	0.043783	1.512 [1.012; 2.260]
DIA_33	-0.012849	0.940405	0.987 [0.705; 1.383]
DIA_34	0.173876	0.365388	1.190 [0.817; 1.734]
DIA_35	-0.279048	0.190565	0.757 [0.498; 1.149]

Continuação da Tabela 1

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
DIA_36	0.183300	0.333594	1.201 [0.828; 1.742]
DIA_37	-0.078778	0.614992	0.924 [0.680; 1.256]
DIA_38	-0.008824	0.962808	0.991 [0.684; 1.436]
DIA_39	0.208294	0.229985	1.232 [0.877; 1.730]
DIA_310	-0.049884	0.790128	0.951 [0.659; 1.374]
DIA_311	-0.442853	0.284280	0.642 [0.286; 1.445]
DIA_312	0.342627	0.131218	1.409 [0.903; 2.198]
DIA_313	-0.010404	0.966268	0.990 [0.611; 1.603]
DIA_314	-0.292455	0.366987	0.746 [0.395; 1.409]
DIA_315	-1.509481	0.033433	0.221 [0.055; 0.888]
DIA_316	0.072550	0.641327	1.075 [0.792; 1.459]
DIA_317	0.001553	0.994282	1.002 [0.655; 1.532]
DIA_318	0.052907	0.711750	1.054 [0.796; 1.396]
DIA_41	0.561281	0.126472	1.753 [0.853; 3.601]
DIA_42	0.258570	0.227209	1.295 [0.851; 1.970]
DIA_43	-0.004285	0.980295	0.996 [0.709; 1.399]
DIA_44	0.385052	0.042528	1.470 [1.013; 2.132]
DIA_45	-0.182641	0.389507	0.833 [0.550; 1.263]
DIA_46	-0.100360	0.640271	0.905 [0.594; 1.378]
DIA_47	0.058677	0.704246	1.060 [0.783; 1.436]
DIA_48	-0.034988	0.862549	0.966 [0.650; 1.435]
DIA_49	0.143872	0.438286	1.155 [0.803; 1.661]
DIA_410	-0.013847	0.943971	0.986 [0.670; 1.451]
DIA_411	-0.174242	0.731910	0.840 [0.310; 2.276]
DIA_412	-0.054279	0.832847	0.947 [0.572; 1.568]
DIA_413	-0.131880	0.649797	0.876 [0.496; 1.549]
DIA_414	-0.274821	0.494922	0.760 [0.345; 1.673]
DIA_415	0.165094	0.805288	1.180 [0.317; 4.383]
DIA_416	0.250732	0.109537	1.285 [0.945; 1.747]
DIA_417	0.381747	0.099153	1.465 [0.931; 2.306]
DIA_418	0.223531	0.127765	1.250 [0.938; 1.667]
DIA_51	0.001155	0.997883	1.001 [0.427; 2.349]
DIA_52	0.292028	0.155026	1.339 [0.895; 2.003]
DIA_53	-0.110346	0.477654	0.896 [0.660; 1.214]
DIA_54	0.459662	0.006527	1.584 [1.137; 2.205]
DIA_55	0.240615	0.195649	1.272 [0.884; 1.831]

Continuação da Tabela 1

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
DIA_56	-0.099231	0.613197	0.906 [0.616; 1.330]
DIA_57	0.061639	0.638561	1.064 [0.822; 1.376]
DIA_58	0.086249	0.621188	1.090 [0.774; 1.535]
DIA_59	0.076525	0.650611	1.080 [0.775; 1.503]
DIA_510	0.174200	0.308529	1.190 [0.851; 1.664]
DIA_511	0.467232	0.365263	1.596 [0.580; 4.387]
DIA_512	0.324532	0.167446	1.383 [0.873; 2.193]
DIA_513	-0.171697	0.547602	0.842 [0.481; 1.474]
DIA_514	-0.375884	0.343921	0.687 [0.315; 1.496]
DIA_515	-0.919329	0.262764	0.399 [0.080; 1.993]
DIA_516	0.048546	0.727558	1.050 [0.799; 1.379]
DIA_517	0.318451	0.163840	1.375 [0.878; 2.153]
DIA_518	-0.060429	0.641531	0.941 [0.730; 1.214]
PROC_13	-1.680979	1.63e-05	0.186 [0.087; 0.400]
PROC_15	-1.259149	0.000953	0.284 [0.135; 0.599]
PROC_16	-1.573337	0.000163	0.207 [0.092; 0.470]
PROC_17	-1.376283	5.25e-06	0.253 [0.140; 0.457]
PROC_18	-1.314264	1.10e-06	0.269 [0.158; 0.456]
PROC_19	-1.330515	0.000235	0.264 [0.130; 0.537]
PROC_110	-0.877907	0.000594	0.416 [0.252; 0.686]
PROC_111	-1.133854	0.000106	0.322 [0.181; 0.571]
PROC_112	-1.589401	0.000172	0.204 [0.089; 0.468]
PROC_113	-1.107544	0.002055	0.330 [0.163; 0.668]
PROC_114	-0.982738	0.017574	0.374 [0.166; 0.842]
PROC_115	-1.997993	2.09e-11	0.136 [0.076; 0.243]
PROC_116	-1.749759	2.73e-08	0.174 [0.094; 0.322]
PROC_117	-1.061723	1.43e-05	0.346 [0.214; 0.559]
PROC_23	-1.920589	0.013367	0.147 [0.032; 0.671]
PROC_25	-0.608792	0.142335	0.544 [0.241; 1.227]
PROC_26	-0.059361	0.895874	0.942 [0.387; 2.292]
PROC_27	-0.213914	0.449230	0.807 [0.464; 1.40]
PROC_28	0.303291	0.135221	1.354 [0.910; 2.016]
PROC_29	-0.003927	0.991250	0.996 [0.494; 2.010]
PROC_210	-0.241340	0.238585	0.786 [0.526; 1.174]
PROC_211	-0.222757	0.377571	0.800 [0.488; 1.313]
PROC_212	-0.512300	0.627123	0.599 [0.076; 4.734]

Continuação da Tabela 1

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
PROC_213	-0.200816	0.639331	0.818 [0.353; 1.895]
PROC_214	-0.191391	0.612724	0.826 [0.394; 1.733]
PROC_215	0.349568	0.246966	1.418 [0.785; 2.563]
PROC_216	-0.448171	0.179472	0.639 [0.332; 1.229]
PROC_217	0.004037	0.977010	1.004 [0.763; 1.321]
PROC_33	1.095254	0.141215	2.990 [0.695; 12.862]
PROC_35	-0.172225	0.674352	0.842 [0.377; 1.880]
PROC_36	0.113376	0.824812	1.120 [0.410; 3.056]
PROC_37	0.058619	0.844375	1.060 [0.591; 1.904]
PROC_38	0.088337	0.683932	1.092 [0.714; 1.671]
PROC_39	-0.200810	0.648748	0.818 [0.345; 1.941]
PROC_310	0.157065	0.495806	1.170 [0.745; 1.839]
PROC_311	-0.067343	0.801546	0.935 [0.553; 1.581]
PROC_312	-12.912521	0.975658	0.000 [0.000; Inf]
PROC_313	0.641464	0.221524	1.899 [0.679; 5.312]
PROC_314	-0.926613	0.107460	0.396 [0.128; 1.223]
PROC_315	0.334808	0.457249	1.398 [0.578; 3.379]
PROC_316	0.416278	0.236806	1.516 [0.761; 3.022]
PROC_317	0.021480	0.876673	1.022 [0.779; 1.340]
PROC_43	0.761542	0.211779	2.142 [0.648; 7.077]
PROC_45	-0.714064	0.151934	0.490 [0.184; 1.301]
PROC_46	-0.586915	0.460603	0.556 [0.117; 2.644]
PROC_47	0.049694	0.872391	1.051 [0.573; 1.927]
PROC_48	0.045505	0.838992	1.047 [0.675; 1.623]
PROC_49	-0.488931	0.298142	0.613 [0.244; 1.541]
PROC_410	-0.109483	0.664603	0.896 [0.546; 1.470]
PROC_411	-0.337370	0.235500	0.714 [0.409; 1.246]
PROC_412	0.378260	0.737359	1.460 [0.160; 13.316]
PROC_413	-12.437994	0.958455	0.000 [0.000; 6.904×10^{197}]
PROC_414	-0.641058	0.261225	0.527 [0.172; 1.612]
PROC_415	-1.318506	0.050800	0.268 [0.071; 1.005]
PROC_416	-0.532359	0.296403	0.587 [0.216; 1.595]
PROC_417	-0.113588	0.403795	0.893 [0.684; 1.165]
PROC_53	0.106206	0.888077	1.112 [0.253; 4.881]
PROC_55	-0.125657	0.819944	0.882 [0.299; 2.602]
PROC_56	0.900284	0.147842	2.460 [0.727; 8.327]

Continuação da Tabela 1

	Coefficiente β	Teste de Wald p-value	OR (IC a 95%)
PROC_57	0.246741	0.426656	1.280 [0.697; 2.352]
PROC_58	-0.320407	0.202412	0.726 [0.443; 1.188]
PROC_59	-13.036607	0.952026	0.000 [0.000; 6.061×10^{178}]
PROC_510	-0.114875	0.680568	0.891 [0.516; 1.540]
PROC_511	-0.345568	0.233459	0.708 [0.401; 1.250]
PROC_512	-12.534899	0.989983	0.000 [0.000; Inf]
PROC_513	-12.760216	0.984826	0.000 [0.000; Inf]
PROC_514	-0.833444	0.155443	0.435 [0.138; 1.372]
PROC_515	0.898984	0.122592	2.457 [0.785; 7.692]
PROC_516	-1.272158	0.220046	0.280 [0.037; 2.140]
PROC_517	0.170526	0.109832	1.186 [0.962; 1.462]

Tabela 2: Coeficientes β , Teste de Wald (p-valor), OR e respectivo IC a 95% relativos ao Modelo 2

	Coeficiente β	Teste Wald p-valor	OR (IC a 95%)
(Intercept)	-0.910271	0.000545	0.402 [0.240; 0.674]
MORF_TUM1	-1.144653	2.78e-07	0.318 [0.206; 0.493]
DSP2	0.320963	0.060769	1.378 [0.986; 1.928]
DSP6	-0.500317	0.037681	0.606 [0.378; 0.972]
DSP7	1.512002	<2e-16	4.536 [3.312; 6.211]
DSP20	0.971994	<2e-16	2.643 [2.165; 3.226]
ICU	-0.030268	0.039920	0.970 [0.943; 0.999]
ADM_TIP2	0.160867	0.024455	1.175 [1.021; 1.351]
ADM_TIP6	-0.915074	0.005484	0.400 [0.210; 0.764]
AGE	-0.007532	4.11e-08	0.992 [0.990; 0.995]
DIA_11	-0.140344	0.801430	0.869 [0.291; 2.594]
DIA_13	-0.262515	0.105597	0.769 [0.560; 1.057]
DIA_14	-0.378216	0.233194	0.685 [0.368; 1.276]
DIA_15	-1.060869	0.000133	0.346 [0.201; 0.597]
DIA_16	-0.855713	1.59e-07	0.425 [0.309; 0.585]
DIA_17	-0.714800	4.36e-11	0.489 [0.396; 0.605]
DIA_18	-0.507564	2.26e-06	0.602 [0.488; 0.743]
DIA_19	-0.376664	0.000477	0.686 [0.555; 0.848]
DIA_110	-0.638722	3.19e-06	0.528 [0.404; 0.691]
DIA_111	-0.057578	0.819405	0.944 [0.576; 1.548]
DIA_112	-0.543075	0.118165	0.581 [0.294; 1.148]
DIA_113	-0.340680	0.114265	0.711 [0.466; 1.086]
DIA_114	-0.787483	0.000738	0.455 [0.288; 0.719]
DIA_115	2.065359	5.60e-08	7.888 [3.744; 16.621]
DIA_116	-0.717733	0.007485	0.488 [0.288; 0.826]
DIA_117	0.192115	0.086531	1.212 [0.973; 1.510]
DIA_118	-0.097715	0.505131	0.907 [0.680; 1.209]
DIA_21	-0.236415	0.608782	0.789 [0.319; 1.952]
DIA_22	0.711450	0.000386	2.037 [1.375; 3.017]
DIA_23	0.035222	0.848695	1.036 [0.721; 1.487]
DIA_24	0.451766	0.024017	1.571 [1.061; 2.326]
DIA_25	0.140816	0.492298	1.151 [0.770; 1.721]
DIA_26	0.205870	0.307265	1.229 [0.828; 1.824]
DIA_27	0.231252	0.169129	1.260 [0.906; 1.752]
DIA_28	0.181025	0.379654	1.198 [0.800; 1.795]
DIA_29	0.124137	0.512355	1.132 [0.781; 1.641]

Continuação da Tabela 2

	Coefficiente β	Teste de Wald p-value	OR (IC a 95%)
DIA_210	0.087647	0.649439	1.092 [0.748; 1.593]
DIA_211	-0.193624	0.641939	0.824 [0.364; 1.864]
DIA_212	0.204876	0.422706	1.227 [0.744; 2.025]
DIA_213	0.164477	0.499732	1.179 [0.731; 1.901]
DIA_214	0.192583	0.485103	1.212 [0.706; 2.082]
DIA_215	-0.891458	0.039789	0.410 [0.175; 0.959]
DIA_216	0.442736	0.007606	1.557 [1.125; 2.155]
DIA_217	0.189651	0.405060	1.209 [0.774; 1.889]
DIA_218	0.218020	0.154018	1.244 [0.922; 1.678]
DIA_31	0.461054	0.143670	1.586 [0.855; 2.942]
DIA_32	0.418722	0.039737	1.520 [1.020; 2.265]
DIA_33	-0.014273	0.933456	0.986 [0.705; 1.378]
DIA_34	0.169426	0.374714	1.185 [0.815; 1.722]
DIA_35	-0.284694	0.179644	0.752 [0.496; 1.140]
DIA_36	0.202981	0.280169	1.225 [0.848; 1.771]
DIA_37	-0.078844	0.613367	0.924 [0.681; 1.255]
DIA_38	0.000281	0.998808	1.000 [0.692; 1.446]
DIA_39	0.205663	0.233302	1.228 [0.876; 1.723]
DIA_310	-0.051543	0.782053	0.950 [0.659; 1.368]
DIA_311	-0.477940	0.242780	0.620 [0.278; 1.383]
DIA_312	0.348347	0.122221	1.417 [0.911; 2.204]
DIA_313	0.019694	0.935716	1.020 [0.632; 1.646]
DIA_314	-0.308608	0.340589	0.734 [0.389; 1.386]
DIA_315	-1.507089	0.032244	0.222 [0.056; 0.880]
DIA_316	0.073385	0.635657	1.076 [0.794; 1.458]
DIA_317	0.002283	0.991539	1.002 [0.657; 1.529]
DIA_318	0.053253	0.708484	1.055 [0.798; 1.394]
DIA_41	0.562774	0.123230	1.756 [0.858; 3.591]
DIA_42	0.248576	0.244033	1.282 [0.844; 1.948]
DIA_43	-0.017993	0.916936	0.982 [0.700; 1.377]
DIA_44	0.387106	0.040040	1.473 [1.018; 2.131]
DIA_45	-0.176725	0.401683	0.838 [0.554; 1.267]
DIA_46	-0.069504	0.743877	0.933 [0.615; 1.415]
DIA_47	0.045384	0.767877	1.046 [0.774; 1.414]
DIA_48	-0.065161	0.745772	0.937 [0.632; 1.389]
DIA_49	0.134670	0.465231	1.144 [0.797; 1.642]

Continuação da Tabela 2

	Coefficiente β	Teste de Wald p-value	OR (IC a 95%)
DIA_410	-0.047924	0.806546	0.953 [0.650; 1.399]
DIA_411	-0.172125	0.725104	0.842 [0.323; 2.197]
DIA_412	-0.065693	0.797332	0.936 [0.567; 1.546]
DIA_413	-0.136051	0.638660	0.873 [0.495; 1.540]
DIA_414	-0.299494	0.454851	0.741 [0.338; 1.626]
DIA_415	0.132661	0.842753	1.142 [0.308; 4.235]
DIA_416	0.249125	0.109107	1.283 [0.946; 1.740]
DIA_417	0.357228	0.119104	1.429 [0.912; 2.240]
DIA_418	0.211334	0.147857	1.235 [0.928; 1.645]
DIA_51	0.024920	0.953996	1.025 [0.440; 2.391]
DIA_52	0.317881	0.119957	1.374 [0.921; 2.051]
DIA_53	-0.085924	0.576762	0.918 [0.679; 1.241]
DIA_54	0.494497	0.003008	1.640 [1.183; 2.273]
DIA_55	0.281426	0.127468	1.325 [0.923; 1.903]
DIA_56	-0.088295	0.650076	0.915 [0.625; 1.341]
DIA_57	0.101936	0.431794	1.107 [0.859; 1.428]
DIA_58	0.142695	0.407612	1.153 [0.823; 1.617]
DIA_59	0.101241	0.544451	1.107 [0.798; 1.535]
DIA_510	0.189569	0.262865	1.209 [0.867; 1.684]
DIA_511	0.182498	0.708488	1.200 [0.461; 3.124]
DIA_512	0.352189	0.129594	1.422 [0.902; 2.243]
DIA_513	-0.151031	0.594929	0.860 [0.493; 1.500]
DIA_514	-0.412054	0.303002	0.662 [0.302; 1.451]
DIA_515	-0.874981	0.302288	0.417 [0.079; 2.198]
DIA_516	0.074444	0.588299	1.077 [0.823; 1.411]
DIA_517	0.318638	0.157080	1.375 [0.885; 2.138]
DIA_518	-0.034889	0.785989	0.966 [0.751; 1.242]
PROC_13	-1.623887	1.59e-05	0.197 [0.094; 0.412]
PROC_15	-1.244383	0.001071	0.288 [0.137; 0.607]
PROC_16	-1.607132	7.75e-05	0.200 [0.090; 0.445]
PROC_17	-1.356980	6.09e-06	0.257 [0.143; 0.463]
PROC_18	-1.348273	4.89e-07	0.260 [0.154; 0.439]
PROC_19	-1.260496	0.000443	0.284 [0.140; 0.573]
PROC_110	-0.905112	0.000379	0.404 [0.246; 0.666]
PROC_111	-1.191624	4.14e-05	0.304 [0.172; 0.537]
PROC_112	-1.621063	0.000122	0.198 [0.086; 0.452]

Continuação da Tabela 2

	Coefficiente β	Teste de Wald p-value	OR (IC a 95%)
PROC_113	-1.125575	0.001560	0.324 [0.162; 0.652]
PROC_114	-1.325443	0.000788	0.266 [0.123; 0.576]
PROC_115	-1.940688	4.03e-11	0.144 [0.081; 0.255]
PROC_116	-1.758042	1.86e-08	0.172 [0.093; 0.318]
PROC_117	-1.053335	1.60e-05	0.349 [0.216; 0.563]
PROC_23	-0.727541	0.125344	0.483 [0.191; 1.225]
PROC_25	-0.567096	0.148789	0.567 [0.263; 1.225]
PROC_26	0.016348	0.968635	1.016 [0.450; 2.296]
PROC_27	-0.184605	0.487151	0.831 [0.494; 1.399]
PROC_28	0.329287	0.055640	1.390 [0.992; 1.947]
PROC_29	0.035933	0.916532	1.037 [0.529; 2.030]
PROC_210	-0.205907	0.255105	0.814 [0.571; 1.160]
PROC_211	-0.225154	0.340046	0.798 [0.503; 1.268]
PROC_212	-0.718262	0.487944	0.488 [0.064; 3.712]
PROC_213	-0.152410	0.710439	0.859 [0.384; 1.920]
PROC_214	-0.226405	0.515831	0.797 [0.403; 1.579]
PROC_215	0.370037	0.163256	1.448 [0.861; 2.436]
PROC_216	-0.338214	0.288226	0.713 [0.382; 1.331]
PROC_217	0.027358	0.808490	1.028 [0.824; 1.282]

Tabela 3: Coeficientes β , Teste de Wald (p-valor), OR e respectivo IC a 95% relativos ao Modelo 3

	Coeficiente β	Teste Wald p-valor	OR (IC a 95%)
(Intercept)	-0.9145442	0.000503	0.401 [0.239; 0.671]
MORF_TUM1	-1.1709220	1.34e-07	0.310 [0.201; 0.479]
DSP2	0.3297684	0.053570	1.391 [0.995; 1.944]
DSP6	-0.5023374	0.036757	0.605 [0.378; 0.970]
DSP7	1.5200551	<2e-16	4.572 [3.342; 6.256]
DSP20	0.9724951	<2e-16	2.645 [2.167; 3.227]
ICU	-0.0303998	0.038841	0.970 [0.942; 0.998]
ADM_TIP2	0.1629935	0.018821	1.177 [1.027; 1.348]
ADM_TIP6	-0.8679916	0.008281	0.420 [0.220; 0.800]
AGE	-0.0072945	7.84e-08	0.993 [0.990; 0.995]
DIA_11	-0.1726750	0.756733	0.841 [0.282; 2.509]
DIA_13	-0.2667276	0.098250	0.766 [0.558; 1.051]
DIA_14	-0.3932050	0.214576	0.675 [0.363; 1.256]
DIA_15	-1.0601123	0.000134	0.346 [0.201; 0.597]
DIA_16	-0.8929431	4.33e-08	0.409 [0.297; 0.564]
DIA_17	-0.7012404	8.39e-11	0.496 [0.401; 0.613]
DIA_18	-0.5163495	1.33e-06	0.597 [0.484; 0.736]
DIA_19	-0.3782804	0.000419	0.685 [0.555; 0.845]
DIA_110	-0.6574076	1.48e-06	0.518 [0.397; 0.677]
DIA_111	-0.0937188	0.702534	0.911 [0.563; 1.473]
DIA_112	-0.5608063	0.106315	0.571 [0.289; 1.127]
DIA_113	-0.3034509	0.155920	0.738 [0.485; 1.123]
DIA_114	-0.7638064	0.001017	0.466 [0.295; 0.735]
DIA_115	2.0631527	5.27e-08	7.871 [3.744; 16.547]
DIA_116	-0.7378075	0.005899	0.478 [0.283; 0.808]
DIA_117	0.1815381	0.103491	1.199 [0.964; 1.492]
DIA_118	-0.0913674	0.528697	0.913 [0.687; 1.213]
DIA_21	-0.2419665	0.599996	0.785 [0.318; 1.939]
DIA_22	0.7057373	0.000406	2.025 [1.370; 2.995]
DIA_23	0.0316380	0.863662	1.032 [0.719; 1.481]
DIA_24	0.4468545	0.025241	1.563 [1.057; 2.312]
DIA_25	0.1320952	0.519051	1.141 [0.764; 1.705]
DIA_26	0.1905058	0.343218	1.210 [0.816; 1.794]
DIA_27	0.2263396	0.177248	1.254 [0.903; 1.742]
DIA_28	0.1727748	0.400477	1.189 [0.795; 1.778]
DIA_29	0.1101012	0.560135	1.116 [0.771; 1.617]

Continuação da Tabela 3

	Coefficiente β	Teste de Wald p-value	OR (IC a 95%)
DIA_210	0.0795604	0.678984	1.083 [0.743; 1.578]
DIA_211	-0.2236647	0.590115	0.800 [0.354; 1.804]
DIA_212	0.1949811	0.444821	1.215 [0.737; 2.004]
DIA_213	0.1637085	0.500837	1.178 [0.731; 1.897]
DIA_214	0.1822632	0.507417	1.200 [0.700; 2.057]
DIA_215	-0.9080245	0.036567	0.403 [0.172; 0.945]
DIA_216	0.4311997	0.009080	1.539 [1.113; 2.128]
DIA_217	0.2003365	0.376447	1.222 [0.784; 1.905]
DIA_218	0.2087325	0.171186	1.232 [0.914; 1.661]
DIA_31	0.4639662	0.141021	1.590 [0.857; 2.950]
DIA_32	0.4238274	0.036785	1.528 [1.026; 2.274]
DIA_33	-0.0007628	0.996433	0.999 [0.715; 1.396]
DIA_34	0.1818467	0.339489	1.199 [0.826; 1.742]
DIA_35	-0.2786432	0.188308	0.757 [0.500; 1.146]
DIA_36	0.1974660	0.292678	1.218 [0.843; 1.760]
DIA_37	-0.0757544	0.626675	0.927 [0.683; 1.258]
DIA_38	0.0067932	0.971113	1.007 [0.697; 1.454]
DIA_39	0.2055906	0.232446	1.228 [0.876; 1.721]
DIA_310	-0.0524819	0.777654	0.949 [0.659; 1.366]
DIA_311	-0.4424996	0.275920	0.642 [0.290; 1.424]
DIA_312	0.3529440	0.116458	1.423 [0.916; 2.211]
DIA_313	0.0325166	0.893902	1.033 [0.641; 1.666]
DIA_314	-0.3145591	0.329967	0.730 [0.388; 1.375]
DIA_315	-1.4823389	0.034431	0.227 [0.057; 0.897]
DIA_316	0.0826054	0.592923	1.086 [0.802; 1.470]
DIA_317	0.0213701	0.920650	1.022 [0.671; 1.556]
DIA_318	0.0614673	0.665460	1.063 [0.805; 1.405]
DIA_41	0.5487211	0.133081	1.731 [0.846; 3.542]
DIA_42	0.2374275	0.264875	1.268 [0.835; 1.925]
DIA_43	-0.0200190	0.907498	0.980 [0.699; 1.374]
DIA_44	0.3953960	0.035800	1.485 [1.027; 2.148]
DIA_45	-0.1761512	0.402686	0.838 [0.555; 1.267]
DIA_46	-0.0791679	0.709361	0.924 [0.609; 1.401]
DIA_47	0.0441829	0.773581	1.045 [0.774; 1.412]
DIA_48	-0.0695637	0.728691	0.933 [0.630; 1.382]
DIA_49	0.1309449	0.476682	1.140 [0.795; 1.635]

Continuação da Tabela 3

	Coefficiente β	Teste de Wald p-value	OR (IC a 95%)
DIA_410	-0.0514810	0.792289	0.950 [0.647; 1.393]
DIA_411	-0.1106904	0.819514	0.895 [0.346; 2.317]
DIA_412	-0.0675003	0.791615	0.935 [0.567; 1.542]
DIA_413	-0.1304360	0.652483	0.878 [0.498; 1.548]
DIA_414	-0.2873840	0.472749	0.750 [0.342; 1.644]
DIA_415	0.1265065	0.849126	1.135 [0.308; 4.178]
DIA_416	0.2510487	0.106032	1.285 [0.948; 1.743]
DIA_417	0.3391715	0.137495	1.404 [0.897; 2.196]
DIA_418	0.2066761	0.156342	1.230 [0.924; 1.636]
DIA_51	0.0154655	0.971425	1.016 [0.436; 2.367]
DIA_52	0.3103474	0.128056	1.364 [0.915; 2.034]
DIA_53	-0.0864560	0.573284	0.917 [0.679; 1.239]
DIA_54	0.4978595	0.002736	1.645 [1.188; 2.279]
DIA_55	0.2748335	0.135866	1.316 [0.917; 1.889]
DIA_56	-0.0990697	0.609968	0.906 [0.619; 1.325]
DIA_57	0.0978451	0.448779	1.103 [0.856; 1.421]
DIA_58	0.1420595	0.408230	1.153 [0.823; 1.614]
DIA_59	0.0841724	0.612663	1.088 [0.785; 1.507]
DIA_510	0.1920263	0.254995	1.212 [0.871; 1.687]
DIA_511	0.1729405	0.721115	1.189 [0.460; 3.072]
DIA_512	0.3496463	0.131180	1.419 [0.901; 2.234]
DIA_513	-0.1398758	0.621909	0.869 [0.499; 1.516]
DIA_514	-0.3822825	0.338725	0.682 [0.312; 1.493]
DIA_515	-0.8982832	0.286530	0.407 [0.078; 2.125]
DIA_516	0.0785302	0.566773	1.082 [0.827; 1.415]
DIA_517	0.3233192	0.149115	1.382 [0.891; 2.144]
DIA_518	-0.0342015	0.789391	0.966 [0.752; 1.242]
PROC_13	-1.9056706	9.97e-09	0.149 [0.078; 0.285]
PROC_15	-1.2266166	0.000920	0.293 [0.142; 0.606]
PROC_16	-1.6673877	6.21e-06	0.189 [0.092; 0.389]
PROC_17	-1.3598729	1.59e-06	0.257 [0.147; 0.447]
PROC_18	-1.2576917	4.89e-07	0.284 [0.174; 0.464]
PROC_19	-1.2324876	0.000410	0.292 [0.147; 0.578]
PROC_110	-0.9473769	7.93e-05	0.388 [0.242; 0.621]
PROC_111	-1.2497272	3.91e-06	0.287 [0.169; 0.487]
PROC_112	-1.6862840	4.40e-05	0.185 [0.082; 0.416]

Continuação da Tabela 3

	Coefficiente β	Teste de Wald p-value	OR (IC a 95%)
PROC_113	-1.1738693	0.000321	0.309 [0.163; 0.586]
PROC_114	-1.3469483	0.000328	0.260 [0.125; 0.542]
PROC_115	-1.8449457	1.97e-11	0.158 [0.092; 0.271]
PROC_116	-1.7895074	1.84e-09	0.167 [0.093; 0.299]
PROC_117	-1.0378683	6.33e-06	0.354 [0.226; 0.556]

Tabela 4: Coeficientes β , Teste de Wald (p-valor), OR e respectivo IC a 95% relativos ao Modelo 4

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
(Intercept)	-0.873249	0.000855	0.418 [0.250; 0.698]
MORF_TUM1	-1.134714	3.13e-07	0.322 [0.208; 0.497]
DSP2	0.331135	0.051862	1.393 [0.997; 1.944]
DSP6	-0.491737	0.040254	0.612 [0.382; 0.978]
DSP7	1.507773	<2e-16	4.517 [3.303; 6.176]
DSP20	0.981129	<2e-16	2.667 [2.190; 3.249]
ICU	-0.026891	0.062619	0.973 [0.946; 1.001]
ADM_TIP2	0.145365	0.044250	1.156 [1.004; 1.332]
ADM_TIP6	-0.871869	0.008083	0.418 [0.219; 0.797]
AGE	-0.007218	2.48e-08	0.993 [0.990; 0.995]
DIA_11	-0.264255	0.637445	0.768 [0.256; 2.304]
DIA_13	-0.370536	0.020549	0.690 [0.505; 0.945]
DIA_14	-0.442096	0.162939	0.643 [0.345; 1.196]
DIA_15	-1.125043	4.52e-05	0.325 [0.189; 0.557]
DIA_16	-0.954330	3.50e-09	0.385 [0.281; 0.529]
DIA_17	-0.781499	1.47e-13	0.458 [0.372; 0.563]
DIA_18	-0.618822	4.04e-09	0.539 [0.438; 0.662]
DIA_19	-0.433617	4.35e-05	0.648 [0.526; 0.798]
DIA_110	-0.727037	7.21e-08	0.483 [0.371; 0.630]
DIA_111	-0.142596	0.556761	0.867 [0.539; 1.395]
DIA_112	-0.622303	0.072013	0.537 [0.272; 1.057]
DIA_113	-0.386191	0.071652	0.680 [0.446; 1.035]
DIA_114	-0.874663	0.000155	0.417 [0.265; 0.656]
DIA_115	1.958210	1.45e-07	7.087 [3.416; 14.703]
DIA_116	-0.780019	0.003502	0.458 [0.272; 0.774]
DIA_117	0.134325	0.220335	1.144 [0.923; 1.418]
DIA_118	-0.133911	0.355105	0.875 [0.659; 1.162]
DIA_21	-0.235025	0.607561	0.791 [0.322; 1.939]
DIA_22	0.690821	0.000540	1.995 [1.349; 2.951]
DIA_23	0.017286	0.925204	1.017 [0.709; 1.460]
DIA_24	0.443596	0.026473	1.558 [1.053; 2.306]
DIA_25	0.096521	0.637479	1.101 [0.737; 1.645]
DIA_26	0.183900	0.359618	1.202 [0.811; 1.781]
DIA_27	0.207160	0.217480	1.230 [0.885; 1.710]
DIA_28	0.167299	0.416012	1.182 [0.790; 1.769]
DIA_29	0.118390	0.530405	1.126 [0.778; 1.629]

Continuação da Tabela 4

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
DIA_210	0.061293	0.749949	1.063 [0.729; 1.550]
DIA_211	-0.213468	0.607739	0.808 [0.358; 1.825]
DIA_212	0.190435	0.455218	1.210 [0.734; 1.994]
DIA_213	0.149335	0.538531	1.161 [0.721; 1.869]
DIA_214	0.140619	0.609995	1.151 [0.671; 1.976]
DIA_215	-0.930669	0.031412	0.394 [0.169; 0.920]
DIA_216	0.420045	0.011024	1.522 [1.101; 2.104]
DIA_217	0.215551	0.339898	1.241 [0.797; 1.931]
DIA_218	0.203617	0.182308	1.226 [0.909; 1.653]
DIA_31	0.581030	0.054914	1.788 [0.988; 3.236]
DIA_32	0.552087	0.003349	1.737 [1.201; 2.512]
DIA_33	0.079761	0.597304	1.083 [0.806; 1.456]
DIA_34	0.304002	0.077408	1.355 [0.967; 1.899]
DIA_35	-0.199847	0.309017	0.819 [0.557; 1.203]
DIA_36	0.281176	0.095126	1.325 [0.952; 1.843]
DIA_37	0.002291	0.986172	1.002 [0.773; 1.299]
DIA_38	0.081543	0.632356	1.085 [0.777; 1.515]
DIA_39	0.347334	0.020629	1.415 [1.055; 1.899]
DIA_310	0.039166	0.814066	1.040 [0.750; 1.441]
DIA_311	-0.383903	0.323331	0.681 [0.318; 1.459]
DIA_312	0.448149	0.030696	1.565 [1.043; 2.350]
DIA_313	0.134647	0.556733	1.144 [0.730; 1.793]
DIA_314	-0.394559	0.202313	0.674 [0.367; 1.236]
DIA_315	-1.688727	0.009614	0.185 [0.051; 0.663]
DIA_316	0.201208	0.121183	1.223 [0.948; 1.577]
DIA_317	0.191949	0.325675	1.212 [0.826; 1.777]
DIA_318	0.183423	0.114003	1.201 [0.957; 1.508]
PROC_13	-1.950388	1.10e-08	0.142 [0.073; 0.278]
PROC_15	-1.256555	0.000680	0.285 [0.138; 0.588]
PROC_16	-1.707409	4.50e-06	0.181 [0.087; 0.376]
PROC_17	-1.375740	1.26e-06	0.253 [0.145; 0.441]
PROC_18	-1.215475	1.26e-06	0.297 [0.181; 0.485]
PROC_19	-1.207444	0.000523	0.299 [0.151; 0.591]
PROC_110	-0.938368	9.46e-05	0.391 [0.244; 0.627]
PROC_111	-1.246097	4.18e-06	0.288 [0.169; 0.489]
PROC_112	-1.694386	3.93e-05	0.184 [0.082; 0.412]

Continuação da Tabela 4

	Coefficiente β	Teste de Wald p-value	OR (IC a 95%)
PROC_113	-1.126597	0.000535	0.324 [0.171; 0.613]
PROC_114	-1.179992	0.001900	0.307 [0.146; 0.647]
PROC_115	-1.866141	1.35e-11	0.155 [0.090; 0.266]
PROC_116	-1.792923	1.77e-09	0.166 [0.093; 0.299]
PROC_117	-1.037422	6.55e-06	0.354 [0.226; 0.556]
PROC_53	0.273633	0.667370	1.315 [0.377; 4.579]
PROC_55	-0.257771	0.645036	0.773 [0.258; 2.314]
PROC_56	0.644066	0.246882	1.904 [0.640; 5.664]
PROC_57	0.255612	0.389147	1.291 [0.722; 2.310]
PROC_58	-0.321693	0.143886	0.725 [0.471; 1.116]
PROC_59	-12.092227	0.929364	0.000 [0.000; 7.274×10^{110}]
PROC_510	-0.104323	0.682482	0.901 [0.547; 1.485]
PROC_511	-0.383259	0.167086	0.682 [0.396; 1.174]
PROC_512	-12.025106	0.984178	0.000 [0.000; Inf]
PROC_513	-11.687778	0.978729	0.000 [0.000; Inf]
PROC_514	-0.691368	0.215500	0.501 [0.168; 1.496]
PROC_515	0.466604	0.346743	1.595 [0.603; 4.215]
PROC_516	-1.352899	0.185868	0.258 [0.035; 1.918]
PROC_517	0.166129	0.021216	1.181 [1.025; 1.360]

Tabela 5: Coeficientes β , Teste de Wald (p-valor), OR e respectivo IC a 95% relativos ao Modelo 5

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
(Intercept)	-0.858670	0.001362	0.424 [0.251; 0.717]
SEXO2	0.011476	0.839188	1.012 [0.905; 1.130]
MORF_TUM1	-1.097979	1.08e-06	0.334 [0.215; 0.519]
DSP2	0.706691	0.518680	2.027 [0.237; 17.337]
DSP6	-0.475900	0.048785	0.621 [0.387; 0.997]
DSP7	1.499676	<2e-16	4.480 [3.262; 6.153]
DSP20	0.952428	<2e-16	2.592 [2.121; 3.168]
PRE_OP	-0.005044	0.514795	0.995 [0.980; 1.010]
ICU	-0.030007	0.047713	0.970 [0.942; 1.000]
ADM_TIP2	0.184220	0.015884	1.202 [1.035; 1.396]
ADM_TIP6	-0.917847	0.005435	0.399 [0.209; 0.763]
AGE	-0.008070	1.37e-08	0.992 [0.989; 0.995]
MOT_TRANF2	-0.369006	0.733342	0.691 [0.083; 5.777]
MOT_TRANF3	-14.360962	0.980314	0.000 [0.000; Inf]
MOT_TRANF4	0.445369	0.753870	1.561 [0.096; 25.265]
DIA_ADM1	-0.941827	0.386100	0.390 [0.046; 3.281]
DIA_ADM3	-0.309330	0.212147	0.734 [0.451; 1.193]
DIA_ADM4	0.014367	0.967092	1.014 [0.513; 2.008]
DIA_ADM5	-0.134344	0.775682	0.874 [0.347; 2.203]
DIA_ADM6	-0.261509	0.286572	0.770 [0.476; 1.245]
DIA_ADM7	-0.099145	0.596633	0.906 [0.627; 1.307]
DIA_ADM8	-0.107610	0.554547	0.898 [0.628; 1.283]
DIA_ADM9	-0.240909	0.175057	0.786 [0.555; 1.113]
DIA_ADM10	-0.259417	0.157948	0.772 [0.538; 1.106]
DIA_ADM11	0.076825	0.852395	1.080 [0.481; 2.426]
DIA_ADM12	-0.449616	0.311311	0.638 [0.267; 1.523]
DIA_ADM13	0.629972	0.086487	1.878 [0.914; 3.858]
DIA_ADM14	-0.048282	0.900098	0.953 [0.448; 2.025]
DIA_ADM15	1.037104	0.110066	2.821 [0.791; 10.066]
DIA_ADM16	-0.284643	0.065943	0.752 [0.555; 1.019]
DIA_ADM17	-0.098008	0.624833	0.907 [0.612; 1.343]
DIA_ADM18	-0.338046	0.075060	0.713 [0.492; 1.035]
DIA_11	0.106948	0.858691	1.113 [0.343; 3.612]
DIA_13	-0.059743	0.798039	0.942 [0.596; 1.489]
DIA_14	-0.358893	0.366303	0.698 [0.321; 1.522]
DIA_15	-0.959825	0.044045	0.383 [0.150; 0.975]

Continuação da Tabela 5

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
DIA_16	-0.680474	0.002306	0.506 [0.327; 0.784]
DIA_17	-0.651840	0.000258	0.521 [0.367; 0.739]
DIA_18	-0.457089	0.009544	0.633 [0.448; 0.895]
DIA_19	-0.208822	0.219704	0.812 [0.581; 1.133]
DIA_110	-0.461172	0.017035	0.631 [0.432; 0.921]
DIA_111	-0.096496	0.829593	0.908 [0.377; 2.186]
DIA_112	-0.269147	0.561727	0.764 [0.308; 1.896]
DIA_113	-0.841239	0.024921	0.431 [0.207; 0.899]
DIA_114	-0.765213	0.032685	0.465 [0.231; 0.939]
DIA_115	1.180752	0.071712	3.257 [0.901; 11.772]
DIA_116	-0.533398	0.065722	0.587 [0.332; 1.035]
DIA_117	0.270483	0.124631	1.311 [0.928; 1.851]
DIA_118	0.097071	0.616816	1.102 [0.753; 1.612]
DIA_21	-0.235360	0.611759	0.790 [0.318; 1.961]
DIA_22	0.736047	0.000303	2.088 [1.400; 3.112]
DIA_23	0.083616	0.653973	1.087 [0.754; 1.567]
DIA_24	0.476431	0.018724	1.610 [1.082; 2.396]
DIA_25	0.154974	0.454515	1.168 [0.778; 1.753]
DIA_26	0.213949	0.294709	1.239 [0.830; 1.848]
DIA_27	0.260622	0.125980	1.298 [0.929; 1.812]
DIA_28	0.216799	0.298170	1.242 [0.826; 1.869]
DIA_29	0.167398	0.382240	1.182 [0.812; 1.721]
DIA_210	0.141070	0.469470	1.152 [0.786; 1.688]
DIA_211	-0.128872	0.759924	0.879 [0.385; 2.009]
DIA_212	0.237331	0.357644	1.268 [0.765; 2.102]
DIA_213	0.164703	0.502462	1.179 [0.729; 1.908]
DIA_214	0.233629	0.399930	1.263 [0.733; 2.176]
DIA_215	-0.887833	0.044757	0.412 [0.173; 0.979]
DIA_216	0.472678	0.004911	1.604 [1.154; 2.230]
DIA_217	0.210653	0.361199	1.234 [0.785; 1.940]
DIA_218	0.255943	0.098638	1.292 [0.953; 1.750]
DIA_31	0.440229	0.164160	1.553 [0.835; 2.888]
DIA_32	0.413397	0.043783	1.512 [1.012; 2.260]
DIA_33	-0.012849	0.940405	0.987 [0.705; 1.383]
DIA_34	0.173876	0.365388	1.190 [0.817; 1.734]
DIA_35	-0.279048	0.190565	0.757 [0.498; 1.149]

Continuação da Tabela 5

	Coefficiente β	Teste de Wald p-value	OR (IC a 95%)
DIA_36	0.183300	0.333594	1.201 [0.828; 1.742]
DIA_37	-0.078778	0.614992	0.924 [0.680; 1.256]
DIA_38	-0.008824	0.962808	0.991 [0.684; 1.436]
DIA_39	0.208294	0.229985	1.232 [0.877; 1.730]
DIA_310	-0.049884	0.790128	0.951 [0.659; 1.374]
DIA_311	-0.442853	0.284280	0.642 [0.286; 1.445]
DIA_312	0.342627	0.131218	1.409 [0.903; 2.198]
DIA_313	-0.010404	0.966268	0.990 [0.611; 1.603]
DIA_314	-0.292455	0.366987	0.746 [0.395; 1.409]
DIA_315	-1.509481	0.033433	0.221 [0.055; 0.888]
DIA_316	0.072550	0.641327	1.075 [0.792; 1.459]
DIA_317	0.001553	0.994282	1.002 [0.655; 1.532]
DIA_318	0.052907	0.711750	1.054 [0.796; 1.396]
DIA_41	0.561281	0.126472	1.753 [0.853; 3.601]
DIA_42	0.258570	0.227209	1.295 [0.851; 1.970]
DIA_43	-0.004285	0.980295	0.996 [0.709; 1.399]
DIA_44	0.385052	0.042528	1.470 [1.013; 2.132]
DIA_45	-0.182641	0.389507	0.833 [0.550; 1.263]
DIA_46	-0.100360	0.640271	0.905 [0.594; 1.378]
DIA_47	0.058677	0.704246	1.060 [0.783; 1.436]
DIA_48	-0.034988	0.862549	0.966 [0.650; 1.435]
DIA_49	0.143872	0.438286	1.155 [0.803; 1.661]
DIA_410	-0.013847	0.943971	0.986 [0.670; 1.451]
DIA_411	-0.174242	0.731910	0.840 [0.310; 2.276]
DIA_412	-0.054279	0.832847	0.947 [0.572; 1.568]
DIA_413	-0.131880	0.649797	0.876 [0.496; 1.549]
DIA_414	-0.274821	0.494922	0.760 [0.345; 1.673]
DIA_415	0.165094	0.805288	1.180 [0.317; 4.383]
DIA_416	0.250732	0.109537	1.285 [0.945; 1.747]
DIA_417	0.381747	0.099153	1.465 [0.931; 2.306]
DIA_418	0.223531	0.127765	1.250 [0.938; 1.667]
DIA_51	0.001155	0.997883	1.001 [0.427; 2.349]
DIA_52	0.292028	0.155026	1.339 [0.895; 2.003]
DIA_53	-0.110346	0.477654	0.896 [0.660; 1.214]
DIA_54	0.459662	0.006527	1.584 [1.137; 2.205]
DIA_55	0.240615	0.195649	1.272 [0.884; 1.831]

Continuação da Tabela 5

	Coefficiente β	Teste de Wald p-value	OR (IC a 95%)
DIA_56	-0.099231	0.613197	0.906 [0.616; 1.330]
DIA_57	0.061639	0.638561	1.064 [0.822; 1.376]
DIA_58	0.086249	0.621188	1.090 [0.774; 1.535]
DIA_59	0.076525	0.650611	1.080 [0.775; 1.503]
DIA_510	0.174200	0.308529	1.190 [0.851; 1.664]
DIA_511	0.467232	0.365263	1.596 [0.580; 4.387]
DIA_512	0.324532	0.167446	1.383 [0.873; 2.193]
DIA_513	-0.171697	0.547602	0.842 [0.481; 1.474]
DIA_514	-0.375884	0.343921	0.687 [0.315; 1.496]
DIA_515	-0.919329	0.262764	0.399 [0.080; 1.993]
DIA_516	0.048546	0.727558	1.050 [0.799; 1.379]
DIA_517	0.318451	0.163840	1.375 [0.878; 2.153]
DIA_518	-0.060429	0.641531	0.941 [0.730; 1.214]
PROC_13	-1.680979	1.63e-05	0.186 [0.087; 0.400]
PROC_15	-1.259149	0.000953	0.284 [0.135; 0.599]
PROC_16	-1.573337	0.000163	0.207 [0.092; 0.470]
PROC_17	-1.376283	5.25e-06	0.253 [0.140; 0.457]
PROC_18	-1.314264	1.10e-06	0.269 [0.158; 0.456]
PROC_19	-1.330515	0.000235	0.264 [0.130; 0.537]
PROC_110	-0.877907	0.000594	0.416 [0.252; 0.686]
PROC_111	-1.133854	0.000106	0.322 [0.181; 0.571]
PROC_112	-1.589401	0.000172	0.204 [0.089; 0.468]
PROC_113	-1.107544	0.002055	0.330 [0.163; 0.668]
PROC_114	-0.982738	0.017574	0.374 [0.166; 0.842]
PROC_115	-1.997993	2.09e-11	0.136 [0.076; 0.243]
PROC_116	-1.749759	2.73e-08	0.174 [0.094; 0.322]
PROC_117	-1.061723	1.43e-05	0.346 [0.214; 0.559]
PROC_23	-1.920589	0.013367	0.147 [0.032; 0.671]
PROC_25	-0.608792	0.142335	0.544 [0.241; 1.227]
PROC_26	-0.059361	0.895874	0.942 [0.387; 2.292]
PROC_27	-0.213914	0.449230	0.807 [0.464; 1.40]
PROC_28	0.303291	0.135221	1.354 [0.910; 2.016]
PROC_29	-0.003927	0.991250	0.996 [0.494; 2.010]
PROC_210	-0.241340	0.238585	0.786 [0.526; 1.174]
PROC_211	-0.222757	0.377571	0.800 [0.488; 1.313]
PROC_212	-0.512300	0.627123	0.599 [0.076; 4.734]

Continuação da Tabela 5

	Coefficiente β	Teste de Wald p-value	OR (IC a 95%)
PROC_213	-0.200816	0.639331	0.818 [0.353; 1.895]
PROC_214	-0.191391	0.612724	0.826 [0.394; 1.733]
PROC_215	0.349568	0.246966	1.418 [0.785; 2.563]
PROC_216	-0.448171	0.179472	0.639 [0.332; 1.229]
PROC_217	0.004037	0.977010	1.004 [0.763; 1.321]
PROC_33	1.095254	0.141215	2.990 [0.695; 12.862]
PROC_35	-0.172225	0.674352	0.842 [0.377; 1.880]
PROC_36	0.113376	0.824812	1.120 [0.410; 3.056]
PROC_37	0.058619	0.844375	1.060 [0.591; 1.904]
PROC_38	0.088337	0.683932	1.092 [0.714; 1.671]
PROC_39	-0.200810	0.648748	0.818 [0.345; 1.941]
PROC_310	0.157065	0.495806	1.170 [0.745; 1.839]
PROC_311	-0.067343	0.801546	0.935 [0.553; 1.581]
PROC_312	-12.912521	0.975658	0.000 [0.000; Inf]
PROC_313	0.641464	0.221524	1.899 [0.679; 5.312]
PROC_314	-0.926613	0.107460	0.396 [0.128; 1.223]
PROC_315	0.334808	0.457249	1.398 [0.578; 3.379]
PROC_316	0.416278	0.236806	1.516 [0.761; 3.022]
PROC_317	0.021480	0.876673	1.022 [0.779; 1.340]
PROC_43	0.761542	0.211779	2.142 [0.648; 7.077]
PROC_45	-0.714064	0.151934	0.490 [0.184; 1.301]
PROC_46	-0.586915	0.460603	0.556 [0.117; 2.644]
PROC_47	0.049694	0.872391	1.051 [0.573; 1.927]
PROC_48	0.045505	0.838992	1.047 [0.675; 1.623]
PROC_49	-0.488931	0.298142	0.613 [0.244; 1.541]
PROC_410	-0.109483	0.664603	0.896 [0.546; 1.470]
PROC_411	-0.337370	0.235500	0.714 [0.409; 1.246]
PROC_412	0.378260	0.737359	1.460 [0.160; 13.316]
PROC_413	-12.437994	0.958455	0.000 [0.000; 6.904×10^{197}]
PROC_414	-0.641058	0.261225	0.527 [0.172; 1.612]
PROC_415	-1.318506	0.050800	0.268 [0.071; 1.005]
PROC_416	-0.532359	0.296403	0.587 [0.216; 1.595]
PROC_417	-0.113588	0.403795	0.893 [0.684; 1.165]
PROC_53	0.106206	0.888077	1.112 [0.253; 4.881]
PROC_55	-0.125657	0.819944	0.882 [0.299; 2.602]
PROC_56	0.900284	0.147842	2.460 [0.727; 8.327]

Continuação da Tabela 5

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
PROC_57	0.246741	0.426656	1.280 [0.697; 2.352]
PROC_58	-0.320407	0.202412	0.726 [0.443; 1.188]
PROC_59	-13.036607	0.952026	0.000 [0.000; 6.061×10^{178}]
PROC_510	-0.114875	0.680568	0.891 [0.516; 1.540]
PROC_511	-0.345568	0.233459	0.708 [0.401; 1.250]
PROC_512	-12.534899	0.989983	0.000 [0.000; Inf]
PROC_513	-12.760216	0.984826	0.000 [0.000; Inf]
PROC_514	-0.833444	0.155443	0.435 [0.138; 1.372]
PROC_515	0.898984	0.122592	2.457 [0.785; 7.692]
PROC_516	-1.272158	0.220046	0.280 [0.037; 2.140]
PROC_517	0.170526	0.109832	1.186 [0.962; 1.462]

Tabela 6: Coeficientes β , Teste de Wald (p-valor), OR e respectivo IC a 95% relativos ao Modelo 6

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
(Intercept)	-0.873249	0.000855	0.418 [0.250; 0.698]
MORF_TUM1	-1.134714	3.13e-07	0.322 [0.208; 0.497]
DSP2	0.331135	0.051862	1.393 [0.997; 1.944]
DSP6	-0.491737	0.040254	0.612 [0.382; 0.978]
DSP7	1.507773	<2e-16	4.517 [3.303; 6.176]
DSP20	0.981129	<2e-16	2.667 [2.190; 3.249]
ICU	-0.026891	0.062619	0.973 [0.946; 1.001]
ADM_TIP2	0.145365	0.044250	1.156 [1.004; 1.332]
ADM_TIP6	-0.871869	0.008083	0.418 [0.219; 0.797]
AGE	-0.007218	2.48e-08	0.993 [0.990; 0.995]
DIA_11	-0.264255	0.637445	0.768 [0.256; 2.304]
DIA_13	-0.370536	0.020549	0.690 [0.505; 0.945]
DIA_14	-0.442096	0.162939	0.643 [0.345; 1.196]
DIA_15	-1.125043	4.52e-05	0.325 [0.189; 0.557]
DIA_16	-0.954330	3.50e-09	0.385 [0.281; 0.529]
DIA_17	-0.781499	1.47e-13	0.458 [0.372; 0.563]
DIA_18	-0.618822	4.04e-09	0.539 [0.438; 0.662]
DIA_19	-0.433617	4.35e-05	0.648 [0.526; 0.798]
DIA_110	-0.727037	7.21e-08	0.483 [0.371; 0.630]
DIA_111	-0.142596	0.556761	0.867 [0.539; 1.395]
DIA_112	-0.622303	0.072013	0.537 [0.272; 1.057]
DIA_113	-0.386191	0.071652	0.680 [0.446; 1.035]
DIA_114	-0.874663	0.000155	0.417 [0.265; 0.656]
DIA_115	1.958210	1.45e-07	7.087 [3.416; 14.703]
DIA_116	-0.780019	0.003502	0.458 [0.272; 0.774]
DIA_117	0.134325	0.220335	1.144 [0.923; 1.418]
DIA_118	-0.133911	0.355105	0.875 [0.659; 1.162]
DIA_21	-0.235025	0.607561	0.791 [0.322; 1.939]
DIA_22	0.690821	0.000540	1.995 [1.349; 2.951]
DIA_23	0.017286	0.925204	1.017 [0.709; 1.460]
DIA_24	0.443596	0.026473	1.558 [1.053; 2.306]
DIA_25	0.096521	0.637479	1.101 [0.737; 1.645]
DIA_26	0.183900	0.359618	1.202 [0.811; 1.781]
DIA_27	0.207160	0.217480	1.230 [0.885; 1.710]
DIA_28	0.167299	0.416012	1.182 [0.790; 1.769]
DIA_29	0.118390	0.530405	1.126 [0.778; 1.629]

Continuação da Tabela 6

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
DIA_210	0.061293	0.749949	1.063 [0.729; 1.550]
DIA_211	-0.213468	0.607739	0.808 [0.358; 1.825]
DIA_212	0.190435	0.455218	1.210 [0.734; 1.994]
DIA_213	0.149335	0.538531	1.161 [0.721; 1.869]
DIA_214	0.140619	0.609995	1.151 [0.671; 1.976]
DIA_215	-0.930669	0.031412	0.394 [0.169; 0.920]
DIA_216	0.420045	0.011024	1.522 [1.101; 2.104]
DIA_217	0.215551	0.339898	1.241 [0.797; 1.931]
DIA_218	0.203617	0.182308	1.226 [0.909; 1.653]
DIA_31	0.581030	0.054914	1.788 [0.988; 3.236]
DIA_32	0.552087	0.003349	1.737 [1.201; 2.512]
DIA_33	0.079761	0.597304	1.083 [0.806; 1.456]
DIA_34	0.304002	0.077408	1.355 [0.967; 1.899]
DIA_35	-0.199847	0.309017	0.819 [0.557; 1.203]
DIA_36	0.281176	0.095126	1.325 [0.952; 1.843]
DIA_37	0.002291	0.986172	1.002 [0.773; 1.299]
DIA_38	0.081543	0.632356	1.085 [0.777; 1.515]
DIA_39	0.347334	0.020629	1.415 [1.055; 1.899]
DIA_310	0.039166	0.814066	1.040 [0.750; 1.441]
DIA_311	-0.383903	0.323331	0.681 [0.318; 1.459]
DIA_312	0.448149	0.030696	1.565 [1.043; 2.350]
DIA_313	0.134647	0.556733	1.144 [0.730; 1.793]
DIA_314	-0.394559	0.202313	0.674 [0.367; 1.236]
DIA_315	-1.688727	0.009614	0.185 [0.051; 0.663]
DIA_316	0.201208	0.121183	1.223 [0.948; 1.577]
DIA_317	0.191949	0.325675	1.212 [0.826; 1.777]
DIA_318	0.183423	0.114003	1.201 [0.957; 1.508]
PROC_13	-1.950388	1.10e-08	0.142 [0.073; 0.278]
PROC_15	-1.256555	0.000680	0.285 [0.138; 0.588]
PROC_16	-1.707409	4.50e-06	0.181 [0.087; 0.376]
PROC_17	-1.375740	1.26e-06	0.253 [0.145; 0.441]
PROC_18	-1.215475	1.26e-06	0.297 [0.181; 0.485]
PROC_19	-1.207444	0.000523	0.299 [0.151; 0.591]
PROC_110	-0.938368	9.46e-05	0.391 [0.244; 0.627]
PROC_111	-1.246097	4.18e-06	0.288 [0.169; 0.489]
PROC_112	-1.694386	3.93e-05	0.184 [0.082; 0.412]

Continuação da Tabela 6

	Coefficiente β	Teste de Wald p-value	OR (IC a 95%)
PROC_113	-1.126597	0.000535	0.324 [0.171; 0.613]
PROC_114	-1.179992	0.001900	0.307 [0.146; 0.647]
PROC_115	-1.866141	1.35e-11	0.155 [0.090; 0.266]
PROC_116	-1.792923	1.77e-09	0.166 [0.093; 0.299]
PROC_117	-1.037422	6.55e-06	0.354 [0.226; 0.556]
PROC_53	0.273633	0.667370	1.315 [0.377; 4.579]
PROC_55	-0.257771	0.645036	0.773 [0.258; 2.314]
PROC_56	0.644066	0.246882	1.904 [0.640; 5.664]
PROC_57	0.255612	0.389147	1.291 [0.722; 2.310]
PROC_58	-0.321693	0.143886	0.725 [0.471; 1.116]
PROC_59	-12.092227	0.929364	0.000 [0.000; 7.274×10^{110}]
PROC_510	-0.104323	0.682482	0.901 [0.547; 1.485]
PROC_511	-0.383259	0.167086	0.682 [0.396; 1.174]
PROC_512	-12.025106	0.984178	0.000 [0.000; Inf]
PROC_513	-11.687778	0.978729	0.000 [0.000; Inf]
PROC_514	-0.691368	0.215500	0.501 [0.168; 1.496]
PROC_515	0.466604	0.346743	1.595 [0.603; 4.215]
PROC_516	-1.352899	0.185868	0.258 [0.035; 1.918]
PROC_517	0.166129	0.021216	1.181 [1.025; 1.360]

Tabela 7: Coeficientes β , Teste de Wald (p-valor), OR e respectivo IC a 95% relativos ao Modelo 7

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
(Intercept)	-0.147743	0.921199	0.863 [0.046; 16.111]
SEXO2	0.011476	0.839188	1.012 [0.905; 1.130]
MORF_TUM1	-1.097979	1.08e-06	0.334 [0.215; 0.519]
PRE_OP	-0.005044	0.514795	0.995 [0.980; 1.010]
ICU	-0.030007	0.047713	0.970 [0.942; 1.000]
AGE	-0.008070	1.37e-08	0.992 [0.989; 0.995]
‘DSP 1‘	-0.952428	<2e-16	0.386 [0.316; 0.471]
‘DSP 2‘	-0.245737	0.823145	0.782 [0.091; 6.748]
‘DSP 6‘	-1.428329	3.72e-08	0.240 [0.144; 0.399]
‘DSP 7‘	0.547247	0.003707	1.728 [1.194; 2.501]
‘DSP 20‘	-	-	-
‘ADM_TIP 1‘	0.917847	0.005435	2.504 [1.311; 4.782]
‘ADM_TIP 2‘	1.102066	0.000945	3.010 [1.566; 5.785]
‘ADM_TIP 6‘	-	-	-
‘MOT_TRANF 0‘	-0.445369	0.753870	0.641 [0.040; 10.367]
‘MOT_TRANF 2‘	-0.814375	0.377485	0.443 [0.073; 2.703]
‘MOT_TRANF 3‘	-14.802756	0.979672	0.000 [0.000; Inf]
‘MOT_TRANF 4‘	-	-	-
‘DIA_ADM 1‘	-0.603781	0.580535	0.547 [0.064; 4.654]
‘DIA_ADM 2‘	0.338046	0.075060	1.402 [0.966; 2.034]
‘DIA_ADM 3‘	0.028716	0.916210	1.029 [0.603; 1.757]
‘DIA_ADM 4‘	0.352413	0.352271	1.422 [0.677; 2.989]
‘DIA_ADM 5‘	0.203702	0.674939	1.226 [0.473; 3.176]
‘DIA_ADM 6‘	0.076537	0.777086	1.080 [0.636; 1.834]
‘DIA_ADM 7‘	0.238901	0.250439	1.270 [0.845; 1.909]
‘DIA_ADM 8‘	0.230437	0.286574	1.259 [0.824; 1.924]
‘DIA_ADM 9‘	0.097137	0.657585	1.102 [0.717; 1.693]
‘DIA_ADM 10‘	0.078629	0.714924	1.082 [0.709; 1.650]
‘DIA_ADM 11‘	0.414871	0.282887	1.514 [0.710; 3.229]
‘DIA_ADM 12‘	-0.111570	0.806533	0.894 [0.366; 2.184]
‘DIA_ADM 13‘	0.968018	0.011208	2.633 [1.246; 5.563]
‘DIA_ADM 14‘	0.289764	0.448362	1.336 [0.632; 2.826]
‘DIA_ADM 15‘	1.375150	0.035759	3.956 [1.096; 14.279]
‘DIA_ADM 16‘	0.053403	0.791578	1.055 [0.710; 1.568]
‘DIA_ADM 17‘	0.240038	0.278616	1.271 [0.823; 1.963]
‘DIA_ADM 18‘	-	-	-

Continuação da Tabela 7

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
‘DIA_1 1’	0.009876	0.987205	1.010 [0.302; 3.377]
‘DIA_1 2’	-0.097071	0.616816	0.907 [0.620; 1.327]
‘DIA_1 3’	-0.156814	0.558930	0.855 [0.505; 1.446]
‘DIA_1 4’	-0.455965	0.284957	0.634 [0.275; 1.462]
‘DIA_1 5’	-1.056897	0.032665	0.348 [0.132; 0.917]
‘DIA_1 6’	-0.777545	0.002920	0.460 [0.275; 0.767]
‘DIA_1 7’	-0.748911	0.000511	0.473 [0.310; 0.721]
‘DIA_1 8’	-0.554161	0.013229	0.575 [0.371; 0.891]
‘DIA_1 9’	-0.305893	0.171569	0.736 [0.475; 1.142]
‘DIA_1 10’	-0.558243	0.016619	0.572 [0.362; 0.904]
‘DIA_1 11’	-0.193567	0.659703	0.824 [0.348; 0.195]
‘DIA_1 12’	-0.366219	0.443882	0.693 [0.272; 1.770]
‘DIA_1 13’	-0.938310	0.016914	0.391 [0.181; 0.845]
‘DIA_1 14’	-0.862284	0.017287	0.422 [0.208; 0.859]
‘DIA_1 15’	1.083681	0.103279	2.956 [0.802; 10.885]
‘DIA_1 16’	-0.630469	0.050096	0.532 [0.283; 1.000]
‘DIA_1 17’	0.173411	0.417085	1.189 [0.782; 1.808]
‘DIA_1 18’	-	-	-
‘DIA_2 0’	-0.255943	0.098638	0.774 [0.571; 1.049]
‘DIA_2 1’	-0.491303	0.267275	0.612 [0.257; 1.457]
‘DIA_2 2’	0.480104	0.002859	1.616 [1.179; 2.216]
‘DIA_2 3’	-0.172326	0.163952	0.842 [0.660; 1.073]
‘DIA_2 4’	0.220488	0.136302	1.247 [0.933; 1.666]
‘DIA_2 5’	-0.100968	0.524187	0.904 [0.663; 1.233]
‘DIA_2 6’	-0.041993	0.784196	0.959 [0.710; 1.295]
‘DIA_2 7’	0.004679	0.961248	1.005 [0.832; 1.213]
‘DIA_2 8’	-0.039143	0.805458	0.962 [0.704; 1.313]
‘DIA_2 9’	-0.088545	0.506916	0.915 [0.705; 1.189]
‘DIA_2 10’	-0.114873	0.403408	0.891 [0.681; 1.167]
‘DIA_2 11’	-0.384814	0.340064	0.681 [0.309; 1.500]
‘DIA_2 12’	-0.018611	0.931915	0.982 [0.640; 1.504]
‘DIA_2 13’	-0.091240	0.654788	0.913 [0.612; 1.362]
‘DIA_2 14’	-0.022314	0.929172	0.978 [0.598; 1.600]
‘DIA_2 15’	-1.143776	0.009355	0.319 [0.134; 0.755]
‘DIA_2 16’	0.216735	0.027246	1.242 [1.025; 1.506]
‘DIA_2 17’	-0.045290	0.810106	0.956 [0.661; 1.383]

Continuação da Tabela 7

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
‘DIA_2 18‘	-	-	-
‘DIA_3 0‘	-0.052907	0.711750	0.948 [0.716; 1.256]
‘DIA_3 1‘	0.387323	0.189304	1.473 [0.826; 2.627]
‘DIA_3 2‘	0.360490	0.042317	1.434 [1.013; 2.031]
‘DIA_3 3‘	-0.065755	0.608542	0.936 [0.728; 1.204]
‘DIA_3 4‘	0.120969	0.437053	1.129 [0.832; 1.531]
‘DIA_3 5‘	-0.331955	0.070385	0.718 [0.501; 1.028]
‘DIA_3 6‘	0.130394	0.392958	1.139 [0.845; 1.537]
‘DIA_3 7‘	-0.131685	0.207522	0.877 [0.714; 1.076]
‘DIA_3 8‘	-0.061730	0.686484	0.940 [0.697; 1.269]
‘DIA_3 9‘	0.155387	0.227980	1.168 [0.907; 1.504]
‘DIA_3 10‘	-0.102790	0.481475	0.902 [0.678; 1.201]
‘DIA_3 11‘	-0.495759	0.220914	0.609 [0.275; 1.347]
‘DIA_3 12‘	0.289720	0.134033	1.336 [0.915; 1.952]
‘DIA_3 13‘	-0.063310	0.770057	0.939 [0.614; 1.435]
‘DIA_3 14‘	-0.345362	0.264567	0.708 [0.386; 1.299]
‘DIA_3 15‘	-1.562387	0.027277	0.210 [0.052; 0.839]
‘DIA_3 16‘	0.019644	0.850972	1.020 [0.831; 1.252]
‘DIA_3 17‘	-0.051353	0.781730	0.950 [0.661; 1.366]
‘DIA_3 18‘	-	-	-
‘DIA_4 0‘	-0.223531	0.127765	0.800 [0.600; 1.066]
‘DIA_4 1‘	0.337750	0.333949	1.402 [0.707; 2.781]
‘DIA_4 2‘	0.035039	0.851984	1.036 [0.717; 1.496]
‘DIA_4 3‘	-0.227815	0.084477	0.796 [0.615; 1.031]
‘DIA_4 4‘	0.161522	0.290684	1.175 [0.871; 1.586]
‘DIA_4 5‘	-0.406171	0.026325	0.666 [0.466; 0.953]
‘DIA_4 6‘	-0.323891	0.076509	0.723 [0.505; 1.035]
‘DIA_4 7‘	-0.164853	0.112327	0.848 [0.692; 1.039]
‘DIA_4 8‘	-0.258519	0.121440	0.772 [0.557; 1.071]
‘DIA_4 9‘	-0.079659	0.579483	0.923 [0.697; 1.224]
‘DIA_4 10‘	-0.237378	0.136642	0.789 [0.577; 1.078]
‘DIA_4 11‘	-0.397773	0.426281	0.672 [0.252; 1.790]
‘DIA_4 12‘	-0.277810	0.225226	0.757 [0.483; 1.187]
‘DIA_4 13‘	-0.355411	0.183441	0.701 [0.415; 1.183]
‘DIA_4 14‘	-0.498351	0.198613	0.608 [0.284; 1.299]
‘DIA_4 15‘	-0.058437	0.930219	0.943 [0.255; 3.489]

Continuação da Tabela 7

	Coefficiente β	Teste de Wald p-value	OR (IC a 95%)
‘DIA_4 16‘	0.027202	0.798901	1.028 [0.834; 1.267]
‘DIA_4 17‘	0.158216	0.433321	1.171 [0.789; 1.740]
‘DIA_4 18‘	-	-	-
‘DIA_5 0‘	0.060429	0.641531	1.062 [0.824; 1.370]
‘DIA_5 1‘	0.061584	0.886233	1.064 [0.457; 2.472]
‘DIA_5 2‘	0.352457	0.073262	1.423 [0.967; 2.092]
‘DIA_5 3‘	-0.049917	0.722002	0.951 [0.723; 1.252]
‘DIA_5 4‘	0.520091	0.000947	1.682 [1.236; 2.290]
‘DIA_5 5‘	0.301044	0.086190	1.351 [0.958; 1.906]
‘DIA_5 6‘	-0.038802	0.834433	0.962 [0.669; 1.384]
‘DIA_5 7‘	0.122068	0.273238	1.130 [0.908; 1.406]
‘DIA_5 8‘	0.146678	0.363093	1.158 [0.844; 1.588]
‘DIA_5 9‘	0.136954	0.380079	1.147 [0.845; 1.557]
‘DIA_5 10‘	0.234629	0.131842	1.264 [0.932; 1.716]
‘DIA_5 11‘	0.527661	0.307759	1.695 [0.615; 4.672]
‘DIA_5 12‘	0.384960	0.086964	1.470 [0.946; 2.284]
‘DIA_5 13‘	-0.111268	0.688310	0.895 [0.519; 1.541]
‘DIA_5 14‘	-0.315455	0.424097	0.729 [0.337; 1.581]
‘DIA_5 15‘	-0.858900	0.296056	0.424 [0.085; 2.122]
‘DIA_5 16‘	0.108975	0.378209	1.115 [0.875; 1.421]
‘DIA_5 17‘	0.378880	0.084641	1.461 [0.950; 2.247]
‘DIA_5 18‘	-	-	-
‘PROC_1 3‘	-1.680979	1.63e-05	0.186 [0.087; 0.400]
‘PROC_1 5‘	-1.259149	0.000953	0.284 [0.135; 0.599]
‘PROC_1 6‘	-1.573337	0.000163	0.207 [0.092; 0.470]
‘PROC_1 7‘	-1.376283	5.25e-06	0.253 [0.140; 0.457]
‘PROC_1 8‘	-1.314264	1.10e-06	0.269 [0.158; 0.456]
‘PROC_1 9‘	-1.330515	0.000235	0.264 [0.130; 0.537]
‘PROC_1 10‘	-0.877907	0.000594	0.416 [0.252; 0.686]
‘PROC_1 11‘	-1.133854	0.000106	0.322 [0.181; 0.571]
‘PROC_1 12‘	-1.589401	0.000172	0.204 [0.089; 0.468]
‘PROC_1 13‘	-1.107544	0.002055	0.330 [0.163; 0.668]
‘PROC_1 14‘	-0.982738	0.017574	0.374 [0.166; 0.842]
‘PROC_1 15‘	-1.997993	2.09e-11	0.136 [0.076; 0.243]
‘PROC_1 16‘	-1.749759	2.73e-08	0.174 [0.094; 0.322]
‘PROC_1 17‘	-1.061723	1.43e-05	0.346 [0.214; 0.559]

Continuação da Tabela 7

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
‘PROC_1 18‘	-	-	-
‘PROC_2 3‘	-1.920589	0.013367	0.147 [0.032; 0.671]
‘PROC_2 5‘	-0.608792	0.142335	0.544 [0.241; 1.227]
‘PROC_2 6‘	-0.059361	0.895874	0.942 [0.387; 2.292]
‘PROC_2 7‘	-0.213914	0.449230	0.807 [0.464; 1.405]
‘PROC_2 8‘	0.303291	0.135221	1.354 [0.910; 2.016]
‘PROC_2 9‘	-0.003927	0.991250	0.996 [0.494; 2.010]
‘PROC_2 10‘	-0.241340	0.238585	0.786 [0.526; 1.174]
‘PROC_2 11‘	-0.222757	0.377571	0.800 [0.488; 1.313]
‘PROC_2 12‘	-0.512300	0.627123	0.599 [0.076; 4.734]
‘PROC_2 13‘	-0.200816	0.639331	0.818 [0.353; 1.895]
‘PROC_2 14‘	-0.191391	0.612724	0.826 [0.394; 1.733]
‘PROC_2 15‘	0.349568	0.246966	1.418 [0.785; 2.563]
‘PROC_2 16‘	-0.448171	0.179472	0.639 [0.332; 1.229]
‘PROC_2 17‘	0.004037	0.977010	1.004 [0.763; 1.321]
‘PROC_2 18‘	-	-	-
‘PROC_3 3‘	1.095254	0.141215	2.990 [0.695; 12.862]
‘PROC_3 5‘	-0.172225	0.674352	0.842 [0.377; 1.880]
‘PROC_3 6‘	0.113376	0.824812	1.120 [0.410; 3.056]
‘PROC_3 7‘	0.058619	0.844375	1.060 [0.591; 1.904]
‘PROC_3 8‘	0.088337	0.683932	1.092 [0.714; 1.671]
‘PROC_3 9‘	-0.200810	0.648748	0.818 [0.345; 1.941]
‘PROC_3 10‘	0.157065	0.495806	1.170 [0.745; 1.839]
‘PROC_3 11‘	-0.067343	0.801546	0.935 [0.553; 1.581]
‘PROC_3 12‘	-12.909562	0.975627	0.000 [0.000; Inf]
‘PROC_3 13‘	0.641464	0.221524	1.899 [0.679; 5.312]
‘PROC_3 14‘	-0.926613	0.107460	0.396 [0.128; 1.223]
‘PROC_3 15‘	0.334808	0.457249	1.398 [0.578; 3.379]
‘PROC_3 16‘	0.416278	0.236806	1.516 [0.761; 3.022]
‘PROC_3 17‘	0.021480	0.876673	1.022 [0.779; 1.340]
‘PROC_3 18‘	-	-	-
‘PROC_4 3‘	0.761542	0.211779	2.142 [0.648; 7.077]
‘PROC_4 5‘	-0.714064	0.151934	0.490 [0.184; 1.301]
‘PROC_4 6‘	-0.586915	0.460603	0.556 [0.117; 2.644]
‘PROC_4 7‘	0.049694	0.872391	1.051 [0.573; 1.927]
‘PROC_4 8‘	0.045505	0.838992	1.047 [0.675; 1.623]

Continuação da Tabela 7

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
‘PROC_4 9‘	-0.488931	0.298142	0.613 [0.244; 1.541]
‘PROC_4 10‘	-0.109483	0.664603	0.896 [0.546; 1.470]
‘PROC_4 11‘	-0.337370	0.235500	0.714 [0.409; 1.246]
‘PROC_4 12‘	0.378260	0.737359	1.460 [0.160; 13.316]
‘PROC_4 13‘	-12.435223	0.958407	0.000 [0.000; 3.621×10^{197}]
‘PROC_4 14‘	-0.641058	0.261225	0.527 [0.172; 1.612]
‘PROC_4 15‘	-1.318506	0.050800	0.268 [0.071; 1.005]
‘PROC_4 16‘	-0.532359	0.296403	0.587 [0.216; 1.595]
‘PROC_4 17‘	-0.113588	0.403795	0.893 [0.684; 1.165]
‘PROC_4 18‘	-	-	-
‘PROC_5 3‘	0.106206	0.888077	1.112 [0.253; 4.881]
‘PROC_5 5‘	-0.125657	0.819944	0.882 [0.299; 2.602]
‘PROC_5 6‘	0.900284	0.147842	2.460 [0.727; 8.327]
‘PROC_5 7‘	0.246741	0.426656	1.280 [0.697; 2.352]
‘PROC_5 8‘	-0.320407	0.202412	0.726 [0.443; 1.188]
‘PROC_5 9‘	-13.033747	0.951968	0.000 [0.000; 3.313×10^{178}]
‘PROC_5 10‘	-0.114875	0.680568	0.891 [0.516; 1.540]
‘PROC_5 11‘	-0.345568	0.233459	0.708 [0.401; 1.250]
‘PROC_5 12‘	-12.531946	0.989971	0.000 [0.000; Inf]
‘PROC_5 13‘	-12.757354	0.984808	0.000 [0.000; Inf]
‘PROC_5 14‘	-0.833444	0.155443	0.435 [0.138; 1.372]
‘PROC_5 15‘	0.898984	0.122592	2.457 [0.785; 7.692]
‘PROC_5 16‘	-1.272158	0.220046	0.280 [0.037; 2.140]
‘PROC_5 17‘	0.170526	0.109832	1.186 [0.962; 1.462]
‘PROC_5 18‘	-	-	-

Tabela 8: Coeficientes β , Teste de Wald (p-valor), OR e respectivo IC a 95% relativos ao Modelo 8

	Coeficiente β	Teste Wald p-valor	OR (IC a 95%)
(Intercept)	-1.294377	0.004287	0.274 [0.113; 0.666]
MORF_TUM1	-1.072085	8.38e-07	0.342 [0.223; 0.524]
ICU	-0.028495	0.049255	0.972 [0.945; 1.000]
AGE	-0.006143	3.09e-07	0.994 [0.992; 0.996]
‘DSP 1‘	-0.333659	0.053083	0.716 [0.511; 1.004]
‘DSP 6‘	-0.821212	0.004939	0.440 [0.248; 0.780]
‘DSP 7‘	1.162398	4.24e-07	3.198 [2.038; 5.017]
‘DSP 20‘	0.722495	0.000179	2.060 [1.411; 3.006]
‘ADM_TIP 1‘	0.952614	0.003573	2.592 [1.366; 4.921]
‘ADM_TIP 2‘	1.080181	0.001029	2.945 [1.545; 5.613]
‘ADM_TIP 6‘	-	-	-
‘MOT_TRANF 4‘	1.040458	0.252007	2.831 [0.477; 16.789]
‘DIA_ADM 13‘	0.740233	0.028008	2.096 [1.083; 4.057]
‘DIA_ADM 15‘	1.979629	2.48e-08	7.240 [3.610; 14.522]
‘DIA_ADM 18‘	-0.286795	0.066926	0.751 [0.552; 1.020]
‘DIA_1 5‘	-0.934905	0.000479	0.393 [0.232; 0.664]
‘DIA_1 6‘	-0.802040	1.41e-07	0.448 [0.333; 0.604]
‘DIA_1 7‘	-0.653258	5.41e-13	0.520 [0.436; 0.621]
‘DIA_1 8‘	-0.480436	9.34e-08	0.619 [0.519; 0.738]
‘DIA_1 10‘	-0.622182	5.51e-07	0.537 [0.421; 0.685]
‘DIA_1 13‘	-0.883197	0.011295	0.413 [0.209; 0.819]
‘DIA_1 14‘	-0.761631	0.000572	0.467 [0.303; 0.720]
‘DIA_1 18‘	0.167278	0.324865	1.182 [0.847; 1.649]
‘DIA_2 2‘	0.576468	0.000156	1.780 [1.320; 2.400]
‘DIA_2 15‘	-1.057006	0.013148	0.347 [0.151; 0.801]
‘DIA_2 16‘	0.317720	0.000241	1.374 [1.160; 1.628]
‘DIA_2 18‘	0.074617	0.239476	1.077 [0.952; 1.220]
‘DIA_3 2‘	0.457751	0.005389	1.581 [1.145; 2.182]
‘DIA_3 5‘	-0.319972	0.065666	0.726 [0.517; 1.021]
‘DIA_3 15‘	-1.472003	0.020465	0.229 [0.066; 0.797]
‘DIA_3 18‘	0.039395	0.573141	1.040 [0.907; 1.193]
‘DIA_4 5‘	-0.203924	0.230530	0.816 [0.584; 1.138]
‘DIA_4 18‘	0.141663	0.055592	1.152 [0.997; 1.332]
‘DIA_5 4‘	0.502050	0.000181	1.652 [1.270; 2.149]
‘DIA_5 18‘	-0.100151	0.243059	0.905 [0.765; 1.070]
‘PROC_1 3‘	-1.495710	5.91e-05	0.224 [0.108; 0.465]

Continuação da Tabela 8

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
‘PROC_1 5‘	-1.218709	0.001181	0.296 [0.142; 0.617]
‘PROC_1 6‘	-1.629825	1.43e-05	0.196 [0.094; 0.409]
‘PROC_1 7‘	-1.310512	8.33e-06	0.270 [0.152; 0.480]
‘PROC_1 8‘	-1.219335	4.04e-06	0.295 [0.176; 0.496]
‘PROC_1 9‘	-1.070769	0.002442	0.343 [0.171; 0.685]
‘PROC_1 10‘	-1.041707	3.06e-05	0.353 [0.216; 0.576]
‘PROC_1 11‘	-1.235332	1.19e-05	0.291 [0.167; 0.505]
‘PROC_1 12‘	-1.718294	3.77e-05	0.179 [0.079; 0.406]
‘PROC_1 13‘	-1.123696	0.000711	0.325 [0.170; 0.623]
‘PROC_1 14‘	-1.553746	1.29e-06	0.211 [0.113; 0.397]
‘PROC_1 15‘	-1.771402	3.35e-10	0.170 [0.098; 0.296]
‘PROC_1 16‘	-1.735612	1.31e-08	0.176 [0.097; 0.321]
‘PROC_1 17‘	-1.039904	1.77e-05	0.353 [0.220; 0.568]
‘PROC_1 18‘	-	-	-
‘PROC_2 3‘	-0.852636	0.068490	0.426 [0.170; 1.067]
‘PROC_2 18‘	-0.033225	0.807508	0.967 [0.740; 1.264]
‘PROC_3 18‘	-0.051372	0.703285	0.950 [0.729; 1.237]
‘PROC_4 18‘	0.093390	0.482617	1.098 [0.846; 1.425]
‘PROC_5 18‘	-0.215586	0.034762	0.806 [0.660; 0.985]

Tabela 9: Coeficientes β , Teste de Wald (p-valor), OR e respectivo IC a 95% relativos ao Modelo 9

	Coeficiente β	Teste Wald p-valor	OR (IC a 95%)
(Intercept)	-1.609559	7.60e-05	0.200 [0.090; 0.444]
MORF_TUM1	-1.072184	7.80e-07	0.342 [0.224; 0.524]
ICU	-0.026397	0.063774	0.974 [0.947; 1.002]
AGE	-0.005770	6.05e-07	0.994 [0.992; 0.997]
‘DSP 6‘	-0.519542	0.029568	0.595 [0.372; 0.950]
‘DSP 7‘	1.479301	<2e-16	4.390 [3.232; 5.962]
‘DSP 20‘	1.045646	<2e-16	2.845 [2.348; 3.448]
‘ADM_TIP 1‘	0.956585	0.003387	2.603 [1.373; 4.935]
‘ADM_TIP 2‘	1.095679	0.000844	2.991 [1.572; 5.692]
‘ADM_TIP 6‘	-	-	-
‘DIA_ADM 13‘	0.776174	0.020825	2.173 [1.125; 4.197]
‘DIA_ADM 15‘	1.997861	1.41e-08	7.373 [3.697; 14.706]
‘DIA_1 5‘	-1.013934	0.000127	0.363 [0.216; 0.609]
‘DIA_1 6‘	-0.811819	9.35e-08	0.444 [0.330; 0.598]
‘DIA_1 7‘	-0.663561	1.70e-13	0.515 [0.432; 0.614]
‘DIA_1 8‘	-0.485642	6.20e-08	0.615 [0.516; 0.734]
‘DIA_1 10‘	-0.612112	7.61e-07	0.542 [0.425; 0.691]
‘DIA_1 13‘	-0.893849	0.009971	0.409 [0.207; 0.807]
‘DIA_1 14‘	-0.779987	0.000370	0.458 [0.298; 0.704]
‘DIA_2 2‘	0.548717	0.000193	1.731 [1.297; 2.310]
‘DIA_2 15‘	-1.141027	0.006262	0.319 [0.141; 0.724]
‘DIA_2 16‘	0.300885	0.000337	1.351 [1.146; 1.593]
‘DIA_3 2‘	0.465403	0.004274	1.593 [1.157; 2.192]
‘DIA_3 15‘	-1.521462	0.015838	0.218 [0.063; 0.752]
‘DIA_5 4‘	0.504631	0.000151	1.656 [1.276; 2.150]
‘PROC_1 3‘	-1.887958	7.92e-09	0.151 [0.080; 0.287]
‘PROC_1 5‘	-1.215659	0.000915	0.297 [0.145; 0.608]
‘PROC_1 6‘	-1.646501	6.25e-06	0.193 [0.094; 0.394]
‘PROC_1 7‘	-1.337319	1.88e-06	0.263 [0.151; 0.455]
‘PROC_1 8‘	-1.249589	4.91e-07	0.287 [0.176; 0.466]
‘PROC_1 9‘	-1.103698	0.001302	0.332 [0.169; 0.650]
‘PROC_1 10‘	-1.061277	7.20e-06	0.346 [0.218; 0.550]
‘PROC_1 11‘	-1.229403	4.45e-06	0.292 [0.173; 0.494]
‘PROC_1 12‘	-1.726822	2.49e-05	0.178 [0.080; 0.397]
‘PROC_1 13‘	-1.149556	0.000351	0.317 [0.169; 0.595]
‘PROC_1 14‘	-1.630983	8.18e-08	0.196 [0.108; 0.355]

Continuação da Tabela 9

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
‘PROC_1 15‘	-1.799200	2.36e-11	0.165 [0.098; 0.280]
‘PROC_1 16‘	-1.764020	1.72e-09	0.171 [0.097; 0.304]
‘PROC_1 17‘	-1.052945	3.85e-06	0.349 [0.223; 0.545]
‘PROC_1 18‘	-	-	-
‘PROC_5 18‘	-0.187812	0.004785	0.829 [0.727; 0.944]

Tabela 10: Coeficientes β , Teste de Wald (p-valor), OR e respectivo IC a 95% relativos ao Modelo 10

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
(Intercept)	-0.617367	0.145775	0.539 [0.235; 1.239]
MORF_TUM1	-1.138762	2.33e-07	0.320 [0.208; 0.493]
ICU	-0.030205	0.039454	0.970 [0.943; 0.999]
AGE	-0.007397	4.71e-09	0.993 [0.990; 0.995]
‘DSP 1‘	-0.983202	<2e-16	0.374 [0.308; 0.455]
‘DSP 6‘	-1.477384	9.27e-09	0.228 [0.138; 0.378]
‘DSP 7‘	0.540302	0.003340	1.717 [1.197; 2.462]
‘ADM_TIP 1‘	0.926697	0.004789	2.526 [1.327; 4.809]
‘ADM_TIP 2‘	1.102296	0.000860	3.011 [1.575; 5.758]
‘MOT_TRANF 2‘	-0.634238	0.000960	0.530 [0.364; 0.773]
‘MOT_TRANF 3‘	-14.758999	0.979607	0.000 [0.000; Inf]
‘DIA_ADM 2‘	0.311261	0.001171	1.365 [1.131; 1.647]
‘DIA_ADM 13‘	0.814472	0.016772	2.258 [1.158; 4.401]
‘DIA_ADM 15‘	1.182530	0.050094	3.263 [1.000; 10.650]
‘DIA_1 5‘	-0.830268	0.002383	0.436 [0.255; 0.745]
‘DIA_1 6‘	-0.630186	0.000107	0.532 [0.387; 0.732]
‘DIA_1 7‘	-0.487101	6.26e-06	0.614 [0.497; 0.759]
‘DIA_1 8‘	-0.307776	0.004154	0.735 [0.596; 0.907]
‘DIA_1 9‘	-0.161475	0.131869	0.851 [0.690; 1.050]
‘DIA_1 10‘	-0.438538	0.001284	0.645 [0.494; 0.842]
‘DIA_1 13‘	-0.753273	0.033618	0.471 [0.235; 0.943]
‘DIA_1 14‘	-0.647440	0.004412	0.523 [0.335; 0.817]
‘DIA_1 15‘	1.141239	0.061752	3.131 [0.945; 10.367]
‘DIA_1 16‘	-0.522378	0.049815	0.593 [0.352; 1.000]
‘DIA_1 17‘	0.413012	0.000113	1.511 [1.225; 1.864]
‘DIA_2 2‘	0.563963	0.000159	1.758 [1.312; 2.355]
‘DIA_2 4‘	0.292751	0.035836	1.340 [1.020; 1.761]
‘DIA_2 15‘	-1.056864	0.012583	0.348 [0.152; 0.797]
‘DIA_2 16‘	0.290647	0.000690	1.337 [1.131; 1.582]
‘DIA_3 2‘	0.387503	0.019474	1.473 [1.064; 2.039]
‘DIA_3 5‘	-0.318411	0.066704	0.727 [0.517; 1.022]
‘DIA_3 9‘	0.185373	0.105524	1.204 [0.962; 1.507]
‘DIA_3 12‘	0.314467	0.085823	1.370 [0.957; 1.961]
‘DIA_3 15‘	-1.645912	0.011364	0.193 [0.054; 0.690]
‘DIA_4 0‘	-0.267679	0.001122	0.765 [0.651; 0.899]
‘DIA_4 3‘	-0.276432	0.021001	0.758 [0.600; 0.959]

Continuação da Tabela 10

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
‘DIA_4 5’	-0.426464	0.013629	0.653 [0.465; 0.916]
‘DIA_4 6’	-0.360640	0.035561	0.697 [0.498; 0.976]
‘DIA_4 7’	-0.207122	0.019270	0.813 [0.683; 0.967]
‘DIA_4 8’	-0.285351	0.066941	0.752 [0.554; 1.020]
‘DIA_4 10’	-0.280874	0.060863	0.755 [0.563; 1.013]
‘DIA_4 12’	-0.308477	0.162342	0.735 [0.477; 1.132]
‘DIA_4 13’	-0.377704	0.146793	0.685 [0.412; 1.142]
‘DIA_4 14’	-0.698626	0.063037	0.497 [0.238; 1.039]
‘DIA_5 2’	0.294931	0.100654	1.343 [0.944; 1.910]
‘DIA_5 4’	0.459037	0.000740	1.583 [1.212; 2.066]
‘DIA_5 5’	0.246307	0.120023	1.279 [0.938; 1.745]
‘DIA_5 12’	0.340096	0.106398	1.405 [0.930; 2.123]
‘DIA_5 17’	0.350610	0.079033	1.420 [0.960; 2.100]
‘PROC_1 3’	-1.712951	5.16e-06	0.180 [0.086; 0.377]
‘PROC_1 5’	-1.257651	0.000671	0.284 [0.138; 0.587]
‘PROC_1 6’	-1.662549	6.71e-06	0.190 [0.092; 0.391]
‘PROC_1 7’	-1.437047	3.84e-07	0.238 [0.136; 0.414]
‘PROC_1 8’	-1.372386	5.84e-08	0.254 [0.154; 0.416]
‘PROC_1 9’	-1.336771	0.000122	0.263 [0.133; 0.520]
‘PROC_1 10’	-1.016182	2.17e-05	0.362 [0.226; 0.579]
‘PROC_1 11’	-1.310389	1.21e-06	0.270 [0.159; 0.458]
‘PROC_1 12’	-1.786383	1.43e-05	0.168 [0.075; 0.375]
‘PROC_1 13’	-1.203946	0.000209	0.300 [0.159; 0.5670]
‘PROC_1 14’	-1.286213	8.93e-05	0.276 [0.145; 0.526]
‘PROC_1 15’	-2.052476	3.21e-13	0.128 [0.074; 0.223]
‘PROC_1 16’	-1.891684	1.75e-10	0.151 [0.084; 0.270]
‘PROC_1 17’	-1.109016	1.35e-06	0.330 [0.210; 0.517]
‘PROC_2 3’	-1.285227	0.021535	0.277 [0.092; 0.828]
‘PROC_2 5’	-0.615624	0.102870	0.540 [0.258; 1.132]
‘PROC_2 8’	0.322072	0.021255	1.380 [1.049; 1.815]
‘PROC_2 15’	0.374201	0.128687	1.454 [0.897; 2.356]
‘PROC_3 14’	-0.926909	0.086716	0.396 [0.137; 1.143]
‘PROC_4 3’	1.223033	0.021650	3.397 [1.196; 9.649]
‘PROC_4 13’	-12.335918	0.959464	0.000 [0.000; 1.728×10^{201}]
‘PROC_5 8’	-0.328794	0.136347	0.720 [0.467; 1.109]
‘PROC_5 9’	-12.982530	0.952239	0.000 [0.000; 7.314×10^{178}]

Continuação da Tabela 10

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
‘PROC_5 11’	-0.426075	0.123996	0.653 [0.379; 1.124]
‘PROC_5 14’	-0.804622	0.150602	0.447 [0.149; 1.340]
‘PROC_5 16’	-1.386562	0.175022	0.250 [0.034; 1.854]
‘PROC_5 17’	0.104275	0.131556	1.110 [0.969; 1.271]

Tabela 11: Coeficientes β , Teste de Wald (p-valor), OR e respectivo IC a 95% relativos ao Modelo 11

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
(Intercept)	-0.147743	0.921199	0.863 [0.046; 16.111]
SEXO2	0.011476	0.839188	1.012 [0.905; 1.130]
MORF_TUM1	-1.097979	1.08e-06	0.334 [0.215; 0.519]
PRE_OP	-0.005044	0.514795	0.995 [0.980; 1.010]
ICU	-0.030007	0.047713	0.970 [0.942; 1.000]
AGE	-0.008070	1.37e-08	0.992 [0.989; 0.995]
‘DSP 1‘	-0.952428	<2e-16	0.386 [0.316; 0.471]
‘DSP 2‘	-0.245737	0.823145	0.782 [0.091; 6.748]
‘DSP 6‘	-1.428329	3.72e-08	0.240 [0.144; 0.399]
‘DSP 7‘	0.547247	0.003707	1.728 [1.194; 2.501]
‘DSP 20‘	-	-	-
‘ADM_TIP 1‘	0.917847	0.005435	2.504 [1.311; 4.782]
‘ADM_TIP 2‘	1.102066	0.000945	3.010 [1.566; 5.785]
‘ADM_TIP 6‘	-	-	-
‘MOT_TRANF 0‘	-0.445369	0.753870	0.641 [0.040; 10.367]
‘MOT_TRANF 2‘	-0.814375	0.377485	0.443 [0.073; 2.703]
‘MOT_TRANF 3‘	-14.802756	0.979672	0.000 [0.000; Inf]
‘MOT_TRANF 4‘	-	-	-
‘DIA_ADM 1‘	-0.603781	0.580535	0.547 [0.064; 4.654]
‘DIA_ADM 2‘	0.338046	0.075060	1.402 [0.966; 2.034]
‘DIA_ADM 3‘	0.028716	0.916210	1.029 [0.603; 1.757]
‘DIA_ADM 4‘	0.352413	0.352271	1.422 [0.677; 2.989]
‘DIA_ADM 5‘	0.203702	0.674939	1.226 [0.473; 3.176]
‘DIA_ADM 6‘	0.076537	0.777086	1.080 [0.636; 1.834]
‘DIA_ADM 7‘	0.238901	0.250439	1.270 [0.845; 1.909]
‘DIA_ADM 8‘	0.230437	0.286574	1.259 [0.824; 1.924]
‘DIA_ADM 9‘	0.097137	0.657585	1.102 [0.717; 1.693]
‘DIA_ADM 10‘	0.078629	0.714924	1.082 [0.709; 1.650]
‘DIA_ADM 11‘	0.414871	0.282887	1.514 [0.710; 3.229]
‘DIA_ADM 12‘	-0.111570	0.806533	0.894 [0.366; 2.184]
‘DIA_ADM 13‘	0.968018	0.011208	2.633 [1.246; 5.563]
‘DIA_ADM 14‘	0.289764	0.448362	1.336 [0.632; 2.826]
‘DIA_ADM 15‘	1.375150	0.035759	3.956 [1.096; 14.279]
‘DIA_ADM 16‘	0.053403	0.791578	1.055 [0.710; 1.568]
‘DIA_ADM 17‘	0.240038	0.278616	1.271 [0.823; 1.963]
‘DIA_ADM 18‘	-	-	-

Continuação da Tabela 11

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
‘DIA_1 1’	0.009876	0.987205	1.010 [0.302; 3.377]
‘DIA_1 2’	-0.097071	0.616816	0.907 [0.620; 1.327]
‘DIA_1 3’	-0.156814	0.558930	0.855 [0.505; 1.446]
‘DIA_1 4’	-0.455965	0.284957	0.634 [0.275; 1.462]
‘DIA_1 5’	-1.056897	0.032665	0.348 [0.132; 0.917]
‘DIA_1 6’	-0.777545	0.002920	0.460 [0.275; 0.767]
‘DIA_1 7’	-0.748911	0.000511	0.473 [0.310; 0.721]
‘DIA_1 8’	-0.554161	0.013229	0.575 [0.371; 0.891]
‘DIA_1 9’	-0.305893	0.171569	0.736 [0.475; 1.142]
‘DIA_1 10’	-0.558243	0.016619	0.572 [0.362; 0.904]
‘DIA_1 11’	-0.193567	0.659703	0.824 [0.348; 0.195]
‘DIA_1 12’	-0.366219	0.443882	0.693 [0.272; 1.770]
‘DIA_1 13’	-0.938310	0.016914	0.391 [0.181; 0.845]
‘DIA_1 14’	-0.862284	0.017287	0.422 [0.208; 0.859]
‘DIA_1 15’	1.083681	0.103279	2.956 [0.802; 10.885]
‘DIA_1 16’	-0.630469	0.050096	0.532 [0.283; 1.000]
‘DIA_1 17’	0.173411	0.417085	1.189 [0.782; 1.808]
‘DIA_1 18’	-	-	-
‘DIA_2 0’	-0.255943	0.098638	0.774 [0.571; 1.049]
‘DIA_2 1’	-0.491303	0.267275	0.612 [0.257; 1.457]
‘DIA_2 2’	0.480104	0.002859	1.616 [1.179; 2.216]
‘DIA_2 3’	-0.172326	0.163952	0.842 [0.660; 1.073]
‘DIA_2 4’	0.220488	0.136302	1.247 [0.933; 1.666]
‘DIA_2 5’	-0.100968	0.524187	0.904 [0.663; 1.233]
‘DIA_2 6’	-0.041993	0.784196	0.959 [0.710; 1.295]
‘DIA_2 7’	0.004679	0.961248	1.005 [0.832; 1.213]
‘DIA_2 8’	-0.039143	0.805458	0.962 [0.704; 1.313]
‘DIA_2 9’	-0.088545	0.506916	0.915 [0.705; 1.189]
‘DIA_2 10’	-0.114873	0.403408	0.891 [0.681; 1.167]
‘DIA_2 11’	-0.384814	0.340064	0.681 [0.309; 1.500]
‘DIA_2 12’	-0.018611	0.931915	0.982 [0.640; 1.504]
‘DIA_2 13’	-0.091240	0.654788	0.913 [0.612; 1.362]
‘DIA_2 14’	-0.022314	0.929172	0.978 [0.598; 1.600]
‘DIA_2 15’	-1.143776	0.009355	0.319 [0.134; 0.755]
‘DIA_2 16’	0.216735	0.027246	1.242 [1.025; 1.506]
‘DIA_2 17’	-0.045290	0.810106	0.956 [0.661; 1.383]

Continuação da Tabela 11

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
‘DIA_2 18‘	-	-	-
‘DIA_3 0‘	-0.052907	0.711750	0.948 [0.716; 1.256]
‘DIA_3 1‘	0.387323	0.189304	1.473 [0.826; 2.627]
‘DIA_3 2‘	0.360490	0.042317	1.434 [1.013; 2.031]
‘DIA_3 3‘	-0.065755	0.608542	0.936 [0.728; 1.204]
‘DIA_3 4‘	0.120969	0.437053	1.129 [0.832; 1.531]
‘DIA_3 5‘	-0.331955	0.070385	0.718 [0.501; 1.028]
‘DIA_3 6‘	0.130394	0.392958	1.139 [0.845; 1.537]
‘DIA_3 7‘	-0.131685	0.207522	0.877 [0.714; 1.076]
‘DIA_3 8‘	-0.061730	0.686484	0.940 [0.697; 1.269]
‘DIA_3 9‘	0.155387	0.227980	1.168 [0.907; 1.504]
‘DIA_3 10‘	-0.102790	0.481475	0.902 [0.678; 1.201]
‘DIA_3 11‘	-0.495759	0.220914	0.609 [0.275; 1.347]
‘DIA_3 12‘	0.289720	0.134033	1.336 [0.915; 1.952]
‘DIA_3 13‘	-0.063310	0.770057	0.939 [0.614; 1.435]
‘DIA_3 14‘	-0.345362	0.264567	0.708 [0.386; 1.299]
‘DIA_3 15‘	-1.562387	0.027277	0.210 [0.052; 0.839]
‘DIA_3 16‘	0.019644	0.850972	1.020 [0.831; 1.252]
‘DIA_3 17‘	-0.051353	0.781730	0.950 [0.661; 1.366]
‘DIA_3 18‘	-	-	-
‘DIA_4 0‘	-0.223531	0.127765	0.800 [0.600; 1.066]
‘DIA_4 1‘	0.337750	0.333949	1.402 [0.707; 2.781]
‘DIA_4 2‘	0.035039	0.851984	1.036 [0.717; 1.496]
‘DIA_4 3‘	-0.227815	0.084477	0.796 [0.615; 1.031]
‘DIA_4 4‘	0.161522	0.290684	1.175 [0.871; 1.586]
‘DIA_4 5‘	-0.406171	0.026325	0.666 [0.466; 0.953]
‘DIA_4 6‘	-0.323891	0.076509	0.723 [0.505; 1.035]
‘DIA_4 7‘	-0.164853	0.112327	0.848 [0.692; 1.039]
‘DIA_4 8‘	-0.258519	0.121440	0.772 [0.557; 1.071]
‘DIA_4 9‘	-0.079659	0.579483	0.923 [0.697; 1.224]
‘DIA_4 10‘	-0.237378	0.136642	0.789 [0.577; 1.078]
‘DIA_4 11‘	-0.397773	0.426281	0.672 [0.252; 1.790]
‘DIA_4 12‘	-0.277810	0.225226	0.757 [0.483; 1.187]
‘DIA_4 13‘	-0.355411	0.183441	0.701 [0.415; 1.183]
‘DIA_4 14‘	-0.498351	0.198613	0.608 [0.284; 1.299]
‘DIA_4 15‘	-0.058437	0.930219	0.943 [0.255; 3.489]

Continuação da Tabela 11

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
‘DIA_4 16’	0.027202	0.798901	1.028 [0.834; 1.267]
‘DIA_4 17’	0.158216	0.433321	1.171 [0.789; 1.740]
‘DIA_4 18’	-	-	-
‘DIA_5 0’	0.060429	0.641531	1.062 [0.824; 1.370]
‘DIA_5 1’	0.061584	0.886233	1.064 [0.457; 2.472]
‘DIA_5 2’	0.352457	0.073262	1.423 [0.967; 2.092]
‘DIA_5 3’	-0.049917	0.722002	0.951 [0.723; 1.252]
‘DIA_5 4’	0.520091	0.000947	1.682 [1.236; 2.290]
‘DIA_5 5’	0.301044	0.086190	1.351 [0.958; 1.906]
‘DIA_5 6’	-0.038802	0.834433	0.962 [0.669; 1.384]
‘DIA_5 7’	0.122068	0.273238	1.130 [0.908; 1.406]
‘DIA_5 8’	0.146678	0.363093	1.158 [0.844; 1.588]
‘DIA_5 9’	0.136954	0.380079	1.147 [0.845; 1.557]
‘DIA_5 10’	0.234629	0.131842	1.264 [0.932; 1.716]
‘DIA_5 11’	0.527661	0.307759	1.695 [0.615; 4.672]
‘DIA_5 12’	0.384960	0.086964	1.470 [0.946; 2.284]
‘DIA_5 13’	-0.111268	0.688310	0.895 [0.519; 1.541]
‘DIA_5 14’	-0.315455	0.424097	0.729 [0.337; 1.581]
‘DIA_5 15’	-0.858900	0.296056	0.424 [0.085; 2.122]
‘DIA_5 16’	0.108975	0.378209	1.115 [0.875; 1.421]
‘DIA_5 17’	0.378880	0.084641	1.461 [0.950; 2.247]
‘DIA_5 18’	-	-	-
‘PROC_1 3’	-1.680979	1.63e-05	0.186 [0.087; 0.400]
‘PROC_1 5’	-1.259149	0.000953	0.284 [0.135; 0.599]
‘PROC_1 6’	-1.573337	0.000163	0.207 [0.092; 0.470]
‘PROC_1 7’	-1.376283	5.25e-06	0.253 [0.140; 0.457]
‘PROC_1 8’	-1.314264	1.10e-06	0.269 [0.158; 0.456]
‘PROC_1 9’	-1.330515	0.000235	0.264 [0.130; 0.537]
‘PROC_1 10’	-0.877907	0.000594	0.416 [0.252; 0.686]
‘PROC_1 11’	-1.133854	0.000106	0.322 [0.181; 0.571]
‘PROC_1 12’	-1.589401	0.000172	0.204 [0.089; 0.468]
‘PROC_1 13’	-1.107544	0.002055	0.330 [0.163; 0.668]
‘PROC_1 14’	-0.982738	0.017574	0.374 [0.166; 0.842]
‘PROC_1 15’	-1.997993	2.09e-11	0.136 [0.076; 0.243]
‘PROC_1 16’	-1.749759	2.73e-08	0.174 [0.094; 0.322]
‘PROC_1 17’	-1.061723	1.43e-05	0.346 [0.214; 0.559]

Continuação da Tabela 11

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
‘PROC_1 18‘	-	-	-
‘PROC_2 3‘	-1.920589	0.013367	0.147 [0.032; 0.671]
‘PROC_2 5‘	-0.608792	0.142335	0.544 [0.241; 1.227]
‘PROC_2 6‘	-0.059361	0.895874	0.942 [0.387; 2.292]
‘PROC_2 7‘	-0.213914	0.449230	0.807 [0.464; 1.405]
‘PROC_2 8‘	0.303291	0.135221	1.354 [0.910; 2.016]
‘PROC_2 9‘	-0.003927	0.991250	0.996 [0.494; 2.010]
‘PROC_2 10‘	-0.241340	0.238585	0.786 [0.526; 1.174]
‘PROC_2 11‘	-0.222757	0.377571	0.800 [0.488; 1.313]
‘PROC_2 12‘	-0.512300	0.627123	0.599 [0.076; 4.734]
‘PROC_2 13‘	-0.200816	0.639331	0.818 [0.353; 1.895]
‘PROC_2 14‘	-0.191391	0.612724	0.826 [0.394; 1.733]
‘PROC_2 15‘	0.349568	0.246966	1.418 [0.785; 2.563]
‘PROC_2 16‘	-0.448171	0.179472	0.639 [0.332; 1.229]
‘PROC_2 17‘	0.004037	0.977010	1.004 [0.763; 1.321]
‘PROC_2 18‘	-	-	-
‘PROC_3 3‘	1.095254	0.141215	2.990 [0.695; 12.862]
‘PROC_3 5‘	-0.172225	0.674352	0.842 [0.377; 1.880]
‘PROC_3 6‘	0.113376	0.824812	1.120 [0.410; 3.056]
‘PROC_3 7‘	0.058619	0.844375	1.060 [0.591; 1.904]
‘PROC_3 8‘	0.088337	0.683932	1.092 [0.714; 1.671]
‘PROC_3 9‘	-0.200810	0.648748	0.818 [0.345; 1.941]
‘PROC_3 10‘	0.157065	0.495806	1.170 [0.745; 1.839]
‘PROC_3 11‘	-0.067343	0.801546	0.935 [0.553; 1.581]
‘PROC_3 12‘	-12.909562	0.975627	0.000 [0.000; Inf]
‘PROC_3 13‘	0.641464	0.221524	1.899 [0.679; 5.312]
‘PROC_3 14‘	-0.926613	0.107460	0.396 [0.128; 1.223]
‘PROC_3 15‘	0.334808	0.457249	1.398 [0.578; 3.379]
‘PROC_3 16‘	0.416278	0.236806	1.516 [0.761; 3.022]
‘PROC_3 17‘	0.021480	0.876673	1.022 [0.779; 1.340]
‘PROC_3 18‘	-	-	-
‘PROC_4 3‘	0.761542	0.211779	2.142 [0.648; 7.077]
‘PROC_4 5‘	-0.714064	0.151934	0.490 [0.184; 1.301]
‘PROC_4 6‘	-0.586915	0.460603	0.556 [0.117; 2.644]
‘PROC_4 7‘	0.049694	0.872391	1.051 [0.573; 1.927]
‘PROC_4 8‘	0.045505	0.838992	1.047 [0.675; 1.623]

Continuação da Tabela 11

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
‘PROC_4 9’	-0.488931	0.298142	0.613 [0.244; 1.541]
‘PROC_4 10’	-0.109483	0.664603	0.896 [0.546; 1.470]
‘PROC_4 11’	-0.337370	0.235500	0.714 [0.409; 1.246]
‘PROC_4 12’	0.378260	0.737359	1.460 [0.160; 13.316]
‘PROC_4 13’	-12.435223	0.958407	0.000 [0.000; 3.621×10^{197}]
‘PROC_4 14’	-0.641058	0.261225	0.527 [0.172; 1.612]
‘PROC_4 15’	-1.318506	0.050800	0.268 [0.071; 1.005]
‘PROC_4 16’	-0.532359	0.296403	0.587 [0.216; 1.595]
‘PROC_4 17’	-0.113588	0.403795	0.893 [0.684; 1.165]
‘PROC_4 18’	-	-	-
‘PROC_5 3’	0.106206	0.888077	1.112 [0.253; 4.881]
‘PROC_5 5’	-0.125657	0.819944	0.882 [0.299; 2.602]
‘PROC_5 6’	0.900284	0.147842	2.460 [0.727; 8.327]
‘PROC_5 7’	0.246741	0.426656	1.280 [0.697; 2.352]
‘PROC_5 8’	-0.320407	0.202412	0.726 [0.443; 1.188]
‘PROC_5 9’	-13.033747	0.951968	0.000 [0.000; 3.313×10^{178}]
‘PROC_5 10’	-0.114875	0.680568	0.891 [0.516; 1.540]
‘PROC_5 11’	-0.345568	0.233459	0.708 [0.401; 1.250]
‘PROC_5 12’	-12.531946	0.989971	0.000 [0.000; Inf]
‘PROC_5 13’	-12.757354	0.984808	0.000 [0.000; Inf]
‘PROC_5 14’	-0.833444	0.155443	0.435 [0.138; 1.372]
‘PROC_5 15’	0.898984	0.122592	2.457 [0.785; 7.692]
‘PROC_5 16’	-1.272158	0.220046	0.280 [0.037; 2.140]
‘PROC_5 17’	0.170526	0.109832	1.186 [0.962; 1.462]
‘PROC_5 18’	-	-	-

Tabela 12: Coeficientes β , Teste de Wald (p-valor), OR e respectivo IC a 95% relativos ao Modelo 12

	Coeficiente β	Teste de Wald p-value	OR (IC a 95%)
(Intercept)	-2.818	3.13e-09	0.060 [0.024; 0.152]
MORF_TUM1	-0.710	0.000413	0.492 [0.332; 0.729]
PRE_OP	-0.011	0.177	0.989 [0.972; 1.005]
ICU	-0.038	0.019	0.963 [0.933; 0.994]
AGE	-0.006	3.65e-07	0.994 [0.992; 0.997]
‘DSP 1’	0.438	0.054	1.550 [0.992; 2.422]
‘DSP 2’	0.806	0.004	2.239 [1.294; 3.873]
‘DSP 7’	2.152	5.86e-16	8.598 [5.106; 14.478]
‘DSP 20’	1.485	1.33e-09	4.416 [2.732; 7.138]
‘ADM_TIP 1’	0.890	0.004	2.436 [1.322; 4.489]
‘ADM_TIP 2’	1.016	0.001	2.763 [1.493; 5.113]
‘DIA_ADM 15’	1.292	0.012	3.640 [1.335; 9.923]
‘DIA_1 15’	1.243	0.018	3.467 [1.232; 9.754]
‘DIA_2 2’	0.740	2.63e-07	2.095 [1.581; 2.777]
‘DIA_2 4’	0.489	0.000359	1.630 [1.246; 2.132]
‘DIA_2 16’	0.331	7.73e-05	1.392 [1.181; 1.640]
‘DIA_3 2’	0.410	0.016	1.506 [1.079; 2.103]
‘DIA_3 4’	0.398	0.002	1.489 [1.151; 1.927]
‘DIA_3 16’	0.130	0.138	1.139 [0.959; 1.353]
‘DIA_4 2’	0.417	0.018	1.517 [1.073; 2.146]
‘DIA_4 4’	0.508	0.000241	1.662 [1.267; 2.180]
‘DIA_4 9’	0.226	0.078	1.253 [0.975; 1.610]
‘DIA_4 16’	0.153	0.107	1.165 [0.967; 1.404]
‘DIA_4 17’	0.446	0.014	1.561 [1.092; 2.232]
‘DIA_4 18’	0.208	0.005	1.231 [1.064; 1.425]
‘DIA_5 2’	0.528	0.004	1.695 [1.182; 2.432]
‘DIA_5 4’	0.663	2.53e-07	1.940 [1.508; 2.496]
‘DIA_5 16’	0.187	0.051	1.206 [1.000; 1.455]
‘DIA_5 17’	0.455	0.027	1.576 [1.053; 2.359]
‘PROC_1 3’	-1.450	0.000186	0.235 [0.110; 0.502]
‘PROC_1 5’	-0.977	0.010	0.377 [0.179; 0.791]
‘PROC_1 6’	-2.114	1.42e-06	0.121 [0.051; 0.285]
‘PROC_1 7’	-1.009	0.000820	0.365 [0.202; 0.658]
‘PROC_1 8’	-1.250	8.32e-06	0.287 [0.165; 0.496]
‘PROC_1 9’	-1.246	0.001	0.288 [0.134; 0.617]
‘PROC_1 10’	-0.822	0.002	0.440 [0.259; 0.747]

Continuação da Tabela 12

	Coefficiente β	Teste de Wald p-value	OR (IC a 95%)
‘PROC_1 11‘	-1.501	9.64e-07	0.223 [0.122; 0.406]
‘PROC_1 12‘	-1.445	0.000501	0.236 [0.104; 0.532]
‘PROC_1 13‘	-0.936	0.006	0.392 [0.200; 0.769]
‘PROC_1 14‘	-0.836	0.010	0.433 [0.230; 0.816]
‘PROC_1 15‘	-1.564	1.53e-07	0.209 [0.117; 0.375]
‘PROC_1 16‘	-1.444	5.40e-06	0.236 [0.127; 0.440]
‘PROC_1 17‘	-0.959	0.000278	0.383 [0.229; 0.643]
‘PROC_2 3‘	-1.408	0.015	0.245 [0.079; 0.758]
‘PROC_2 6‘	0.941	0.007	2.562 [1.291; 5.087]
‘PROC_2 18‘	-0.203	0.086	0.816 [0.648; 1.029]
‘PROC_4 3‘	1.028	0.098	2.796 [0.828; 9.440]
‘PROC_4 17‘	0.186	0.065	1.204 [0.989; 1.466]
‘PROC_4 18‘	0.285	0.014	1.329 [1.059; 1.668]

Apêndice B
Código utilizado no Software RStudio

Inicialmente é feito o carregamento das bibliotecas necessárias e a leitura dos dados:

```
library(readr)
library(lsr)
library(ryouready)
library(pROC)
library(ResourceSelection)
library(psc1)
library(MASS)
library(car)
library(reshape2)

dt <- read_delim("data", ";",
  escape_double = FALSE, col_types = cols(AGE = col_number(),
  BIRTH_WGT = col_number(), DISTANCIA = col_number(),
  ICU = col_number(), PRE_OP = col_number(),
  TOTDIAS = col_number()), trim_ws = TRUE)
```

Elimina-se as observações duplicadas:

```
dt<-as.data.frame(dt)
#encontrar observações duplicadas na variável NUM_EPISOD:
vec<-which(duplicated(dt$NUM_EPISOD))
dt<-dt[-vec,] #eliminar essas observações da base de dados
```

Elimina-se a observação codificada com 3 na variável SEXO:

```
dt<-dt[-140150,]
```

Isola-se apenas o MODULO codificado com INT que corresponde ao módulo de internamento:

```
#encontrar as observações correspondentes ao módulo de internamento:
vt<-which(dt$MODULO=="INT")
dt<-dt[vt,] #isolar essas observações
```

Eliminação de variáveis com valores omissos superiores a 80% e não relevantes para o estudo em questão:

```
dt<-dt[,-c(2,5,7:70,77:111,118:131,150:152,154,155,157:160,163,166)]
dt<-dt[,-c(1,16,23,24,32,42,45:48)]
dt<-dt[,-c(33,37,38)]
```

Eliminação das observações sem diagnóstico principal atribuído:

```
#encontrar as observações com valores omissos da variável DDX1:
idx_out<-which(is.na(dt$DDX1)==TRUE)
dt<-dt[-idx_out,] #eliminar essas observações
```

Recodificação das variáveis relativas aos diagnósticos e aos procedimentos:

```
#abrir o ficheiro com os códigos de diagnóstico:
diagn <- read_delim("Diag2.txt","\t", escape_double = FALSE,
  col_types = cols(D10 = col_number(), D11 = col_number(),
  D12 = col_number(), D13 = col_number(), D14 = col_number(),
  D15 = col_number(), D16 = col_number(), D17 = col_number(),
  D2 = col_number(), D3 = col_number(), D4 = col_number(),
  D5 = col_number(), D6 = col_number(), D7 = col_number(),
  D8 = col_number(), D9 = col_number()), trim_ws = TRUE)

#matriz de zeros, com 6 colunas e com número de linhas igual ao da base de dados:
```

```

temp2<-as.data.frame(matrix(0,nrow=nrow(dt),ncol=6))
for (l in 4:9){
  for (i in 1:nrow(dt)){
    if(!is.na(dt[i,l])){ #se a variável está preenchida
      for (j in 1:ncol(diagn)){
        var10<-c(diagn[,j])
        if (dt[i,l]%in%var10[[1]]==TRUE){
          temp2[i,1-3]=j #preenche na matriz com o código do grupo de diagnóstico
          next #próxima observação
        }
      }
    }
  }
}

#como a matriz é uma matriz de zeros, se a observação não estiver preenchida fica
#automaticamente com o código 0, correspondente a diagnóstico não efetuado
#alterar o nome das novas variáveis:
names(temp2)<- c("DIA_ADM", "DIA_1", "DIA_2","DIA_3","DIA_4","DIA_5")
write.csv(temp2, file = "diag.csv") #gravar o ficheiro

#abrir o ficheiro com os códigos dos procedimentos:
proc <- read_delim("proc.txt", "\t", escape_double = FALSE,
  col_types = cols(`0` = col_character(), `1` = col_character(),
    `2` = col_character(), `3` = col_character()))

#matriz de 18, com 5 colunas e com número de linhas igual ao da base de dados:
temp3<-as.data.frame(matrix(18,nrow=nrow(dt),ncol=5))
for (l in 10:14){
  for (i in 1:nrow(dt)){
    if(!is.na(dt[i,l])){ #se a variável está preenchida
      for (j in 1:ncol(proc)){
        var50<-c(proc[,j])
        if (dt[i,l]%in%var50[[1]]==TRUE){
          temp3[i,1-5]=j-1
          #preenche na matriz com o código do grupo do procedimento aplicado
          next #próxima observação
        }
      }
    }
  }
}

#como a matriz é uma matriz com 18 em todas as entradas, se a observação não estiver
#preenchida fica automaticamente com o código 18, correspondente a procedimento não efetuado
names(temp3)<- c("PROC_1", "PROC_2","PROC_3","PROC_4","PROC_5")
write.csv(temp3,file="proced.csv")

```

Eliminação das variáveis ADM_DIAGN, DDX1, DDX2, DDX3, DDX4, DDX5, SRG1, SRG2, SRG3, SRG4, SRG5:

```
dt<-dt[,-c(4:14)]
```

Eliminação das variáveis NUMERO, NUM_EPISD, FIN_RESP, DSP_GDH, GDH_HCFA16 e GCD_HCFA16:

```
dt<-dt[,-c(1,16)]
dt<-dt[,-c(2,11,13,15)]
```

Recodificação da variável MORF_TUM:

```
dt$MORF_TUM<-as.numeric(dt$MORF_TUM)
for (i in 1:nrow(dt)){
  if(is.na(dt$MORF_TUM[i])){ #se a variável não está preenchida
    dt$MORF_TUM[i]<-0 #colocar 0
  }
  else {
    dt$MORF_TUM[i]<-1 #caso contrário, colocar 1
  }
}
dt$MORF_TUM<-as.factor(dt$MORF_TUM)
```

Base de dados final:

```
#juntar à base da dados as variáveis construídas relativas aos diagnósticos e aos
#procedimentos:
dtf<-as.data.frame(cbind(dt,temp2,temp3))
```

Cálculo das medidas de associação:

```
#cria uma matriz vazia com 1 coluna e 22 linhas:
corr2<-as.data.frame(matrix(ncol=1,nrow=22))
for (j in 1:22){ #para cada variável
  if (is.numeric(dtf[,j])){ #se é numérica
    #preenche o vetor com o respectivo coeficiente eta:
    corr2[j,1]<-eta(dtf[,11],dtf[,j])
  }
  else if (is.factor(dtf[,j])){ #se é categórica
    #preenche o vetor com o respectivo coeficiente Cramer:
    corr2[j,1]<-cramersV(dtf[,j],dtf[,11])
  }
}
corr2<-round(corr2,4) #arredonda a matriz com 4 casas decimais
write.csv(corr2, file = "Cor_Tent.csv") #grava a matriz para posterior análise
```

Construção de modelos de regressão logística simples:

```
dtf<-as.data.frame(dt)

##Fatores
#tabela de contingência:
addmargins(acast(dtf, SEXO ~ RINT))
#modelo de regressão logística simples:
mod1=glm(RINT~SEXO, data=dtf, family="binomial"(logit))
summary(mod1) #sumário do modelo
OR1=exp(mod1$coefficients); #valor da razão de chances
#intervalo de confiança para o parâmetro estimado:
ICbeta1=confint.default(mod1,level=0.95);
ICOR1=exp(ICbeta1); #intervalo de confiança para a razão de chances
#agregação da razão de chances e do respetivo intervalo de
#confiança com 3 casas decimais:
round((cbind(OR1, ICOR1)),3)
```

```

addmargins(acast(dtf, MORF_TUM ~ RINT))
mod2=glm(RINT~MORF_TUM, data=dtf, family="binomial"(logit))
summary(mod2)
OR2=exp(mod2$coefficients);
ICbeta2=confint.default(mod2,level=0.95);
ICOR2=exp(ICbeta2);
round((cbind(OR2, ICOR2)),3)

addmargins(acast(dtf, DSP ~ RINT))
mod3=glm(RINT~DSP, data=dtf, family="binomial"(logit))
summary(mod3)
OR3=exp(mod3$coefficients);
ICbeta3=confint.default(mod3,level=0.95);
ICOR3=exp(ICbeta3);
round((cbind(OR3, ICOR3)),3)

addmargins(acast(dtf, ADM_TIP ~ RINT))
mod4=glm(RINT~ADM_TIP, data=dtf, family="binomial"(logit))
summary(mod4)
OR4=exp(mod4$coefficients);
ICbeta4=confint.default(mod4,level=0.95);
ICOR4=exp(ICbeta4);
round((cbind(OR4, ICOR4)),3)

addmargins(acast(dtf, MOT_TRANF ~ RINT))
mod5=glm(RINT~MOT_TRANF, data=dtf, family="binomial"(logit))
summary(mod5)
OR5=exp(mod5$coefficients);
ICbeta5=confint.default(mod5,level=0.95);
ICOR5=exp(ICbeta5);
round((cbind(OR5, ICOR5)),3)

addmargins(acast(dtf, DIA_ADM ~ RINT))
mod7=glm(RINT~relevel(DIA_ADM, ref=2), data=dtf, family="binomial"(logit))
summary(mod7)
OR7=exp(mod7$coefficients);
ICbeta7=confint.default(mod7,level=0.95);
ICOR7=exp(ICbeta7);
round((cbind(OR7, ICOR7)),3)

addmargins(acast(dtf, DIA_1 ~ RINT))
mod8=glm(RINT~relevel(DIA_1, ref=2), data=dtf, family="binomial"(logit))
summary(mod8)
OR8=exp(mod8$coefficients);
ICbeta8=confint.default(mod8,level=0.95);
ICOR8=exp(ICbeta8);
round((cbind(OR8, ICOR8)),3)

addmargins(acast(dtf, DIA_2 ~ RINT))
mod9=glm(RINT~DIA_2, data=dtf, family="binomial"(logit))
summary(mod9)
OR9=exp(mod9$coefficients);
ICbeta9=confint.default(mod9,level=0.95);

```

```

ICOR9=exp(ICbeta9);
round((cbind(OR9, ICOR9)),3)

addmargins(acast(dtf, DIA_3 ~ RINT))
mod10=glm(RINT~DIA_3, data=dtf, family="binomial"(logit))
summary(mod10)
OR10=exp(mod10$coefficients);
ICbeta10=confint.default(mod10,level=0.95);
ICOR10=exp(ICbeta10);
round((cbind(OR10, ICOR10)),3)

addmargins(acast(dtf, DIA_4 ~ RINT))
mod11=glm(RINT~DIA_4, data=dtf, family="binomial"(logit))
summary(mod11)
OR11=exp(mod11$coefficients);
ICbeta11=confint.default(mod11,level=0.95);
ICOR11=exp(ICbeta11);
round((cbind(OR11, ICOR11)),3)

addmargins(acast(dtf, DIA_5 ~ RINT))
mod12=glm(RINT~DIA_5, data=dtf, family="binomial"(logit))
summary(mod12)
OR12=exp(mod12$coefficients);
ICbeta12=confint.default(mod12,level=0.95);
ICOR12=exp(ICbeta12);
round((cbind(OR12, ICOR12)),3)

addmargins(acast(dtf, PROC_1 ~ RINT))
mod13=glm(RINT~relevel(PROC_1, ref=15), data=dtf, family="binomial"(logit))
summary(mod13)
OR13=exp(mod13$coefficients);
ICbeta13=confint.default(mod13,level=0.95);
ICOR13=exp(ICbeta13);
round((cbind(OR13, ICOR13)),3)
round((cbind(mod13$coefficients,OR13, ICOR13)),3)

addmargins(acast(dtf, PROC_2 ~ RINT))
mod14=glm(RINT~relevel(PROC_2, ref=15), data=dtf, family="binomial"(logit))
summary(mod14)
OR14=exp(mod14$coefficients);
ICbeta14=confint.default(mod14,level=0.95);
ICOR14=exp(ICbeta14);
round((cbind(OR14, ICOR14)),3)
round((cbind(mod14$coefficients,OR14, ICOR14)),3)

addmargins(acast(dtf, PROC_3 ~ RINT))
mod15=glm(RINT~relevel(PROC_3, ref=15), data=dtf, family="binomial"(logit))
summary(mod15)
OR15=exp(mod15$coefficients);
ICbeta15=confint.default(mod15,level=0.95);
ICOR15=exp(ICbeta15);
round((cbind(OR15, ICOR15)),3)
round((cbind(mod15$coefficients,OR15, ICOR15)),3)

```



```

addmargins(acast(dtf, PROC_4 ~ RINT))
mod16=glm(RINT~relevel(PROC_4, ref=15), data=dtf, family="binomial"(logit))
summary(mod16)
OR16=exp(mod16$coefficients);
ICbeta16=confint.default(mod16,level=0.95);
ICOR16=exp(ICbeta16);
round((cbind(OR16, ICOR16)),3)
round((cbind(mod16$coefficients,OR16, ICOR16)),3)

addmargins(acast(dtf, PROC_5 ~ RINT))
mod17=glm(RINT~relevel(PROC_5, ref=15), data=dtf, family="binomial"(logit))
summary(mod17)
OR17=exp(mod17$coefficients);
ICbeta17=confint.default(mod17,level=0.95);
ICOR17=exp(ICbeta17);
round((cbind(OR17, ICOR17)),3)
round((cbind(mod17$coefficients,OR17, ICOR17)),3)

#Quantitativas:
mod18=glm(RINT~BIRTH_WGT, data=dtf, family="binomial"(logit))
summary(mod18)
OR18=exp(mod18$coefficients);
ICbeta18=confint.default(mod18,level=0.95);
ICOR18=exp(ICbeta18);
round((cbind(OR18, ICOR18)),4)

mod19=glm(RINT~PRE_OP, data=dtf, family="binomial"(logit))
summary(mod19)
OR19=exp(mod19$coefficients);
ICbeta19=confint.default(mod19,level=0.95);
ICOR19=exp(ICbeta19);
round((cbind(OR19, ICOR19)),3)

mod20=glm(RINT~ICU, data=dtf, family="binomial"(logit))
summary(mod20)
OR20=exp(mod20$coefficients);
ICbeta20=confint.default(mod20,level=0.95);
ICOR20=exp(ICbeta20);
round((cbind(OR20, ICOR20)),3)

mod21=glm(RINT~TOTDIAS, data=dtf, family="binomial"(logit))
summary(mod21)
OR21=exp(mod21$coefficients);
ICbeta21=confint.default(mod21,level=0.95);
ICOR21=exp(ICbeta21);
round((cbind(OR21, ICOR21)),3)

mod22=glm(RINT~AGE, data=dtf, family="binomial"(logit))
summary(mod22)
OR22=exp(mod22$coefficients);
ICbeta22=confint.default(mod22,level=0.95);
ICOR22=exp(ICbeta22);
round((cbind(OR22, ICOR22)),3)

```

Transformação das variáveis com mais de duas categorias em variáveis binárias:

```
trans_to_bin <- function(df){
  tmp2=as.data.frame(numeric(nrow(df)));
  tmp2=tmp2[, -1]
  rem=numeric(0)
  for(i in 1:ncol(df)){
    if(is.factor(df[,i])){
      if(nlevels(df[,i])>2){
        rem<-c(rem,i)
        tmp<-as.data.frame(model.matrix(~factor(df[,i])-1))
        c<-as.data.frame(strsplit(colnames(tmp), "[ ]"))
        a<-colnames(df)[i]
        f<-c(as.matrix(c[2,]))
        colnames(tmp)<-paste(a,f)
        tmp2<-merge(tmp2,tmp,by="row.names",all.x=TRUE,all.y = TRUE)
        tmp2[,1]<-as.integer(tmp2[,1])
        tmp2<-tmp2[order(tmp2$Row.names),]
        rownames(tmp2)<-tmp2[,1]
        tmp2<-tmp2[, -1]
      }
    }
  }
  dfr<-df[, -rem]
  dft<-cbind(dfr,tmp2)
  return(dft)
}
```

Esta função foi desenvolvida por João Oliveira, aluno da Universidade de Aveiro do Mestrado de Matemática e Aplicações.

Cosntrução dos modelos de regressão logística múltiplos:

```
##### MOLDELOS MULTIVARIADOS #####
##### VARIÁVEIS ORIGINAIS #####

# Da análise bivariada sai: BIRTH_WGT, TOTDIAS
dr10<-dtf[, -c(4,8)]
## Modelo 1
set.seed(5678) #seed para se ter sempre a mesma amostra
ind10 = sample(nrow(dr10), floor(nrow(dr10) * 0.7))
drtr10 = dr10[ind10,] #conjunto de treino
drte10 = dr10[-ind10,] #conjunto de teste

#modelo de regressão logística com alteração das classes de referência:
fit1=glm(RINT~SEXO+MORF_TUM+DSP+PRE_OP+ICU+ADM_TIP+AGE+MOT_TRANF+
  relevel(DIA_ADM, ref = 2)+ relevel(DIA_1, ref = 2)+DIA_2+DIA_3+
  DIA_4+DIA_5+relevel(PROC_1, ref = 15) + relevel(PROC_2, ref = 15)+
  relevel(PROC_3, ref = 15) + relevel(PROC_4, ref = 15)+
  relevel(PROC_5, ref = 15),data=drtr10, family="binomial"(logit))
summary(fit1)
ORfit1=exp(fit1$coefficients);
ICbetafit1=confint.default(fit1,level=0.95);
ICORfit1=exp(ICbetafit1);
round((cbind(ORfit1, ICORfit1)),3)
# Curva ROC
```

```

predict.Mod1<-predict(fit1, drte10, type="response") #previsão de valores
roc_full_resolution1 <- roc(drte10$RINT, predict.Mod1) #construção da curva ROC
plot(roc_full_resolution1, print.auc=TRUE,
     ylab="Sensibilidade", xlab="Especificidade")
coords(roc_full_resolution1, "best") #ponto de corte
# Matriz de confusao
premod1<-predict(fit1, drte10, type="response")
pre1<-ifelse(premod1>0.1000111,1,0) #aplicar o ponto de corte
aMF1<-caret::confusionMatrix(pre1, drte10$RINT) #matriz de confusão
# Teste de Hosmer e Lemeshow
drtr10$RINT<-as.numeric(drtr10$RINT)
drtr10$RINT<-ifelse(drtr10$RINT=="1",0,1)
hoslem.test(drtr10$RINT, fitted(fit1), g=10)
#AIC e BIC:
AIC(fit1)
BIC(fit1)

## Modelo 2
dr10<-dtf[,-c(4,8)]
dr20<-dr10[,-c(1,4,8,10,18,19,20)]
set.seed(5678)
ind20 = sample(nrow(dr20),floor(nrow(dr20) * 0.7))
drtr20 = dr20[ind20,] #conjunto de treino
drte20 = dr20[-ind20,] #conjunto de teste

fit2=glm(RINT~MORF_TUM+DSP+ICU+ADM_TIP+AGE+relevel(DIA_1, ref = 2)+DIA_2+
        DIA_3+DIA_4+DIA_5+relevel(PROC_1, ref = 15) + relevel(PROC_2, ref = 15),
        data=drtr20, family="binomial"(logit))
summary(fit2)
ORfit2=exp(fit2$coefficients);
ICbetafit2=confint.default(fit2,level=0.95);
ICORfit2=exp(ICbetafit2);
round((cbind(ORfit2, ICORfit2)),3)
# Curva ROC
predict.Mod2<-predict(fit2, drte20, type="response")
roc_full_resolution2 <- roc(drte20$RINT, predict.Mod2)
plot(roc_full_resolution2, print.auc=TRUE,
     ylab="Sensibilidade", xlab="Especificidade")
coords(roc_full_resolution2, "best")
# Matriz de confusao
premod2<-predict(fit2, drte20, type="response")
pre2<-ifelse(premod2>=0.09288369,1,0)
aMF2<-caret::confusionMatrix(pre2, drte20$RINT)
# Teste de Hosmer e Lemeshow
drtr20$RINT<-as.numeric(drtr20$RINT)
drtr20$RINT<-ifelse(drtr20$RINT=="1",0,1)
hoslem.test(drtr20$RINT, fitted(fit2), g=10)
# AIC e BIC
AIC(fit2)
BIC(fit2)

## Modelo 3
dr10<-dtf[,-c(4,8)]

```

```

dr30<-dr10[,-c(1,4,8,10,17,18,19,20)]
set.seed(5678)
ind30 = sample(nrow(dr30),floor(nrow(dr30) * 0.7))
drtr30 = dr30[ind30,] #conjunto de treino
drte30 = dr30[-ind30,] #conjunto de teste

fit3=glm(RINT~MORF_TUM+DSP+ICU+ADM_TIP+AGE+relevel(DIA_1, ref = 2)+DIA_2+
        DIA_3+DIA_4+DIA_5+relevel(PROC_1, ref = 15),
        data=drtr20, family="binomial"(logit))
summary(fit3)
ORfit3=exp(fit3$coefficients);
ICbetafit3=confint.default(fit3,level=0.95);
ICORfit3=exp(ICbetafit3);
round((cbind(ORfit3, ICORfit3)),3)
# Curva ROC
predict.Mod3<-predict(fit3, drte30, type="response")
roc_full_resolution3 <- roc(drte30$RINT, predict.Mod3)
plot(roc_full_resolution3, print.auc=TRUE,
     ylab="Sensibilidade", xlab="Especificidade")
coords(roc_full_resolution3, "best")
# Matriz de confusao
premod3<-predict(fit3, drte30, type="response")
pre3<-ifelse(premod3>=0.1004262,1,0)
aMF3<-caret::confusionMatrix(pre3, drte30$RINT)
# Teste de Hosmer e Lemeshow
drtr30$RINT<-as.numeric(drtr30$RINT)
drtr30$RINT<-ifelse(drtr30$RINT=="1",0,1)
hoslem.test(drtr30$RINT, fitted(fit3), g=10)
# AIC e BIC
AIC(fit3)
BIC(fit3)

##### MOLDELOS STEPWISE #####
##### VARIÁVEIS ORIGINAIS #####

#BACWARD
set.seed(5678)
ind10 = sample(nrow(dr10),floor(nrow(dr10) * 0.7))
drtr10 = dr10[ind10,] #conjunto de treino
drte10 = dr10[-ind10,] #conjunto de teste

stepb <- stepAIC(fit1, direction="backward") #stepwise com direção backward
summary(stepb)
ORfitsb=exp(stepb$coefficients);
ICbetafitsb=confint.default(stepb,level=0.95);
ICORfitsb=exp(ICbetafitsb);
round((cbind(stepb$coefficients,ORfitsb, ICORfitsb)),3)
#Curva ROC
predict.Modb <- predict(stepb, drte10, type="response")
roc_full_resolutionb <- roc(drte10$RINT, predict.Modb)
plot(roc_full_resolutionb, print.auc=TRUE,
     ylab="Sensibilidade", xlab="Especificidade")
coords(roc_full_resolutionb, "best")

```

```

#Matriz de Confusão
preb<-ifelse(predict.Modb>=0.1125579,1,0)
aMFb<-caret::confusionMatrix(preb, drte10$RINT)
# Teste de Hosmer e Lemeshow
drtr10$RINT<-as.numeric(drtr10$RINT)
drtr10$RINT<-ifelse(drtr10$RINT=="1",0,1)
hoslem.test(drtr10$RINT, fitted(stepb), g=10)
# AIC e BIC
AIC(stepb)
BIC(stepb)

#FORWARD
set.seed(5678)
ind10 = sample(nrow(dr10),floor(nrow(dr10) * 0.7))
drtr10 = dr10[ind10,] #conjunto de treino
drte10 = dr10[-ind10,] #conjunto de teste

stepf <- stepAIC(fit1, direction="forward") #stepwise com direção forward
summary(stepf)
ORfitsf=exp(stepf$coefficients);
ICbetafitsf=confint.default(stepf,level=0.95);
ICORfitsf=exp(ICbetafitsf);
round((cbind(ORfitsf, ICORfitsf)),3)
#Curva ROC
predict.Modf <- predict(stepf, drte10, type="response")
roc_full_resolutionf <- roc(drte10$RINT, predict.Modf)
plot(roc_full_resolutionf, print.auc=TRUE,
     ylab="Sensibilidade", xlab="Especificidade")
coords(roc_full_resolutionf, "best")
#Matriz de Confusão
pref<-ifelse(predict.Modf>=0.1000111,1,0)
aMFf<-caret::confusionMatrix(pref, drte10$RINT)
# Teste de Hosmer e Lemeshow
drtr10$RINT<-as.numeric(drtr10$RINT)
drtr10$RINT<-ifelse(drtr10$RINT=="1",0,1)
hoslem.test(drtr10$RINT, fitted(stepf), g=10)
# AIC e BIC
AIC(stepf)
BIC(stepf)

#BOTH
set.seed(5678)
ind10 = sample(nrow(dr10),floor(nrow(dr10) * 0.7))
drtr10 = dr10[ind10,] #conjunto de treino
drte10 = dr10[-ind10,] #conjunto de teste

stepbo <- stepAIC(fit1, direction="both") #stepwise com direção both
summary(stepbo)
ORfitsbo=exp(stepbo$coefficients);
ICbetafitsbo=confint.default(stepbo,level=0.95);
ICORfitsbo=exp(ICbetafitsbo);
round((cbind(ORfitsbo, ICORfitsbo)),3)
#Curva ROC

```

```

predict.Modbo <- predict(stepbo, drte10, type="response")
roc_full_resolutionbo <- roc(drte10$RINT, predict.Modbo)
plot(roc_full_resolutionbo, print.auc=TRUE,
     ylab="Sensibilidade", xlab="Especificidade")
coords(roc_full_resolutionbo, "best")
#Matriz de Confusão
prebo<-ifelse(predict.Modbo>=0.1125579,1,0)
aMFbo<-caret::confusionMatrix(prebo, drte10$RINT)
# Teste de Hosmer e Lemeshow
drtr10$RINT<-as.numeric(drtr10$RINT)
drtr10$RINT<-ifelse(drtr10$RINT=="1",0,1)
hoslem.test(drtr10$RINT, fitted(stepbo), g=10)
# AIC e BIC
AIC(stepbo)
BIC(stepbo)

##### MOLDELOS MULTIVARIADOS #####
##### VARIÁVEIS BINÁRIAS #####

# Da análise bivariada sai: BIRTH_WGT, TOTDIAS
dtf1<-dtf[,-c(4,8)]
dr<-trans_to_bin(dtf1)
dr1<-dr
set.seed(5678)
ind11 = sample(nrow(dr1),floor(nrow(dr1) * 0.7))
drtr11 = dr1[ind11,] #conjunto de treino
drte11= dr1[-ind11,] #conjunto de teste

fit11=glm(RINT~., data=drtr11, family="binomial"(logit))
summary(fit11)
ORfit11=exp(fit11$coefficients);
ICbetafit11=confint.default(fit11,level=0.95);
ICORfit11=exp(ICbetafit11);
round((cbind(ORfit11, ICORfit11)),3)
# Curva ROC
predict.Mod11<-predict(fit11, drte11, type="response")
roc_full_resolution11 <- roc(drte11$RINT, predict.Mod11)
plot(roc_full_resolution11, print.auc=TRUE,
     ylab="Sensibilidade", xlab="Especificidade")
coords(roc_full_resolution11, "best")
# Matriz de confusao
pre11<-ifelse(predict.Mod11>=0.1000111,1,0)
aMF11<-caret::confusionMatrix(pre11, drte11$RINT)
# Teste de Hosmer e Lemeshow
drtr11$RINT<-as.numeric(drtr11$RINT)
drtr11$RINT<-ifelse(drtr11$RINT=="1",0,1)
hoslem.test(drtr11$RINT, fitted(fit11), g=10)
# AIC e BIC
AIC(fit11)
BIC(fit11)

##Modelo 2:
dr1<-dr

```

```

dr2<-dr1[,c(2,4:7,9:14,18,31,33,36,41:44,46,49,50,54,57,70,71,73,76,79,89,
           92,98,111,116,130:146,160,175,190,205)]
set.seed(5678)
ind22 = sample(nrow(dr2),floor(nrow(dr2) * 0.7))
drtr22 = dr2[ind22,] #conjunto de treino
drte22 = dr2[-ind22,] #conjunto de teste

fit22=glm(RINT~., data=drtr22, family="binomial"(logit))
summary(fit22)
ORfit22=exp(fit22$coefficients);
ICbetafit22=confint.default(fit22,level=0.95);
ICORfit22=exp(ICbetafit22);
round((cbind(ORfit22, ICORfit22)),3)
# Curva ROC
predict.Mod22<-predict(fit22, drte22, type="response")
roc_full_resolution22 <- roc(drte22$RINT, predict.Mod22)
plot(roc_full_resolution22, print.auc=TRUE,
     ylab="Sensibilidade", xlab="Especificidade")
coords(roc_full_resolution22, "best")
# Matriz de confusao
pre22<-ifelse(predict.Mod22>=0.08500905,1,0)
aMF22<-caret::confusionMatrix(pre22, drte22$RINT)
# Teste de Hosmer e Lemeshow
drtr22$RINT<-as.numeric(drtr22$RINT)
drtr22$RINT<-ifelse(drtr22$RINT=="1",0,1)
hoslem.test(drtr22$RINT, fitted(fit22), g=10)
# AIC e BIC
AIC(fit22)
BIC(fit22)

##Modelo 3:
dr1<-dr
dr2<-dr1[,c(2,4:7,9:14,18,31,33,36,41:44,46,49,50,54,57,70,71,73,76,79,89,
           92,98,111,116,130:146,160,175,190,205)]
dr3<-dr2[,c(5,12,15,23,27,29,31:33,35,51:54)]
set.seed(5678)
ind33 = sample(nrow(dr3),floor(nrow(dr3) * 0.7))
drtr33 = dr3[ind33,] #conjunto de treino
drte33 = dr3[-ind33,] #conjunto de teste

fit33=glm(RINT~., data=drtr33, family="binomial"(logit))
summary(fit33)
ORfit33=exp(fit33$coefficients);
ICbetafit33=confint.default(fit33,level=0.95);
ICORfit33=exp(ICbetafit33);
round((cbind(ORfit33, ICORfit33)),3)
# Curva ROC
predict.Mod33<-predict(fit33, drte33, type="response")
roc_full_resolution33 <- roc(drte33$RINT, predict.Mod33)
plot(roc_full_resolution33, print.auc=TRUE,
     ylab="Sensibilidade", xlab="Especificidade")
coords(roc_full_resolution33, "best")
# Matriz de confusao

```

```

pre33<-ifelse(predict.Mod33>=0.09014899,1,0)
aMF33<-caret::confusionMatrix(pre33, drte33$RINT)
# Teste de Hosmer e Lemeshow
drtr33$RINT<-as.numeric(drtr33$RINT)
drtr33$RINT<-ifelse(drtr33$RINT=="1",0,1)
hoslem.test(drtr33$RINT, fitted(fit33), g=10)
# AIC e BIC
AIC(fit33)
BIC(fit33)

##### MOLDELOS STEPWISE #####
##### VARIÁVEIS BINÁRIAS #####

#BACKWARD:
set.seed(5678)
dr1<-dr
ind11 = sample(nrow(dr1),floor(nrow(dr1) * 0.7))
drtr1 = dr1[ind11,] #conjunto de treino
drte1= dr1[-ind11,]

stepbB <- stepAIC(fit11, direction="backward")
summary(stepbB)
ORfitsbB=exp(stepbB$coefficients);
ICbetafitsbB=confint.default(stepbB,level=0.95);
ICORfitsbB=exp(ICbetafitsbB);
round((cbind(ORfitsbB, ICORfitsbB)),3)
#Curva ROC
predict.ModbB<-predict(stepbB, drte1, type="response")
roc_full_resolutionbB <- roc(drte1$RINT, predict.ModbB)
plot(roc_full_resolutionbB, print.auc=TRUE,
     ylab="Sensibilidade", xlab="Especificidade")
coords(roc_full_resolutionbB, "best")
#Matriz de Confusão
prebB<-ifelse(predict.ModbB>=0.1023365,1,0)
aMFbB<-caret::confusionMatrix(prebB, drte1$RINT)
# Teste de Hosmer e Lemeshow
drtr1$RINT<-as.numeric(drtr1$RINT)
drtr1$RINT<-ifelse(drtr1$RINT=="1",0,1)
hoslem.test(drtr1$RINT, fitted(stepbB), g=10)
# AIC e BIC
AIC(stepbB)
BIC(stepbB)

#FORWARD
dr1<-dr
set.seed(5678)
ind11 = sample(nrow(dr1),floor(nrow(dr1) * 0.7))
drtr1 = dr1[ind11,] #conjunto de treino
drte1= dr1[-ind11,]

stepfB <- stepAIC(fit11, direction="forward")
summary(stepfB)
ORfitsfB=exp(stepfB$coefficients);

```



```

ICbetafitsfB=confint.default(stepfB,level=0.95);
ICORfitsfB=exp(ICbetafitsfB);
round((cbind(ORfitsfB, ICORfitsfB)),3)
#Curva ROC
predict.ModfB<-predict(stepfB, drte1, type="response")
roc_full_resolutionfB <- roc(drte1$RINT, predict.ModfB)
plot(roc_full_resolutionfB, print.auc=TRUE,
     ylab="Sensibilidade", xlab="Especificidade")
coords(roc_full_resolutionfB, "best")
#Matriz de Confusão
prefB<-ifelse(predict.ModfB>=0.1000111,1,0)
aMFfB<-caret::confusionMatrix(prefB, drte1$RINT)
# Teste de Hosmer e Lemeshow
drtr1$RINT<-as.numeric(drtr1$RINT)
drtr1$RINT<-ifelse(drtr1$RINT=="1",0,1)
hoslem.test(drtr1$RINT, fitted(stepfB), g=10)
# AIC e BIC
AIC(stepfB)
BIC(stepfB)

#BOTH
dr1<-dr
set.seed(5678)
ind11 = sample(nrow(dr1),floor(nrow(dr1) * 0.7))
drtr1 = dr1[ind11,] #conjunto de treino
drte1= dr1[-ind11,]

stepBo <- stepAIC(fit11, direction="both")
summary(stepBo)
ORfitsBo=exp(stepBo$coefficients);
ICbetafitsBo=confint.default(stepBo,level=0.95);
ICORfitsBo=exp(ICbetafitsBo);
round((cbind(ORfitsBo, ICORfitsBo)),3)
#Curva ROC
predict.ModboB <-predict(stepBo, drte1, type="response")
roc_full_resolutionboB <- roc(drte1$RINT, predict.ModboB)
plot(roc_full_resolutionboB, print.auc=TRUE,
     ylab="Sensibilidade", xlab="Especificidade")
coords(roc_full_resolutionboB, "best")
#Matriz de Confusão
preboB<-ifelse(predict.ModboB>=0.1134052,1,0)
aMFboB<-caret::confusionMatrix(preboB, drte1$RINT)
# Teste de Hosmer e Lemeshow
drtr1$RINT<-as.numeric(drtr1$RINT)
drtr1$RINT<-ifelse(drtr1$RINT=="1",0,1)
hoslem.test(drtr1$RINT, fitted(stepBo), g=10)
# AIC e BIC
AIC(stepBo)
BIC(stepBo)

```